Wireless Communications

Andrea J. Goldsmith

Second Edition

Contents

1 Overview of Wireless Communications **1**

1.1 History of Wireless Communications . . . 1
1.1.1 Origins of Radio Technology . . . 1
1.1.2 From Analog to Digital . . . 2
1.1.3 Evolution of Wireless Systems and Standards . . . 3
1.2 Wireless Network Paradigms . . . 7
1.2.1 Wireless Local Area Networks . . . 8
1.2.2 Cellular Networks . . . 9
1.2.3 Satellite Systems . . . 12
1.2.4 Fixed Wireless Access . . . 13
1.2.5 Ad Hoc Wireless Networks . . . 14
1.3 Wireless Spectrum . . . 15
1.3.1 Licensed and Unlicensed Spectral Bands . . . 16
1.3.2 Underlay Systems and Cognitive Radios . . . 17
1.3.3 Spectrum Allocations . . . 18
1.4 Standards . . . 19
1.5 Wireless Vision . . . 20
1.6 Technical Challenges . . . 21

2 Path Loss, Shadowing, and Multipath **27**

2.1 Radio Wave Propagation . . . 28
2.2 Transmit and Receive Signal Models . . . 29
2.3 Free-Space Path Loss . . . 31
2.4 Two-Ray Multipath Model . . . 33
2.5 Path Loss Exponent Models . . . 36
2.5.1 Single-Slope . . . 36
2.5.2 Multi-Slope . . . 38
2.6 Shadowing . . . 39
2.7 Combined Path Loss and Shadowing . . . 43
2.7.1 Single-Slope Path Loss with Log-Normal Shadowing . . . 43
2.7.2 Outage Probability . . . 44
2.7.3 Cell Coverage Area and Percentage . . . 44
2.8 General Ray Tracing . . . 48
2.8.1 Multi-Ray Reflections . . . 49
2.8.2 Diffraction . . . 50
2.8.3 Scattering . . . 52

2.8.4 Multipath Model with Reflection, Diffraction, and Scattering . . . 53
2.8.5 Local Mean Received Power . . . 53
2.9 Directional Antennas, Antenna Arrays, and Intelligent Surfaces . . . 53
2.10 Measurement-Based Propagation Models . . . 55
2.10.1 Okumura Model . . . 56
2.10.2 Hata Model . . . 57
2.10.3 Cellular System Channel Models . . . 57
2.10.4 Wi-Fi Channel Models . . . 59
2.10.5 Millimeter Wave Models . . . 60
2.10.6 Indoor Attenuation Models . . . 62

3 Statistical Multipath Channel Models 76
3.1 Time-Varying Channel Impulse Response . . . 77
3.2 Narrowband Fading Model . . . 82
3.2.1 Autocorrelation, Cross-Correlation, and Power Spectral Density . . . 83
3.2.2 Envelope and Power Distributions for Narrowband Fading . . . 89
3.2.3 Level Crossing Rate and Average Fade Duration . . . 93
3.2.4 Block-Fading and Finite-State Markov Fading . . . 95
3.3 Wideband Fading Model . . . 96
3.3.1 Autocorrelation and Scattering Function . . . 97
3.3.2 Power Delay Profile . . . 99
3.3.3 Coherence Bandwidth . . . 101
3.3.4 Doppler Power Spectrum and Channel Coherence Time . . . 103
3.3.5 Transforms for Autocorrelation and Scattering Functions . . . 104
3.4 Discrete-Time Model . . . 105
3.5 MIMO Channel Models . . . 106

4 Capacity of Wireless Channels 115
4.1 Capacity in AWGN . . . 116
4.2 Capacity of Flat Fading Channels . . . 118
4.2.1 Channel and System Model . . . 118
4.2.2 Channel Distribution Information Known . . . 119
4.2.3 Channel Side Information at Receiver . . . 119
4.2.4 Channel Side Information at Transmitter and Receiver . . . 122
4.2.5 Capacity with Receiver Diversity . . . 128
4.2.6 Capacity Comparisons . . . 129
4.3 Capacity of Frequency-Selective Fading Channels . . . 131
4.3.1 Time-Invariant Channels . . . 131
4.3.2 Time-Varying Channels . . . 133

5 Digital Modulation and Detection 143
5.1 Signal Space Analysis . . . 144
5.1.1 Signal and System Model . . . 144
5.1.2 Geometric Representation of Signals . . . 145
5.1.3 Receiver Structure and Sufficient Statistics . . . 148
5.1.4 Decision Regions and the Maximum Likelihood Decision Criterion . . . 151
5.1.5 Error Probability and the Union Bound . . . 153

5.2 Passband Modulation Principles . . . 158
5.3 Amplitude and Phase Modulation . . . 158
5.3.1 Pulse Amplitude Modulation (MPAM) . . . 160
5.3.2 Phase-Shift Keying (MPSK) . . . 162
5.3.3 Quadrature Amplitude Modulation (MQAM) . . . 164
5.3.4 Differential Modulation . . . 165
5.3.5 Constellation Shaping . . . 168
5.3.6 Quadrature Offset . . . 168
5.4 Frequency Modulation . . . 168
5.4.1 Frequency-Shift Keying (FSK) and Minimum-Shift Keying (MSK) . . . 169
5.4.2 Continuous-Phase FSK (CPFSK) . . . 170
5.4.3 Noncoherent Detection of FSK . . . 171
5.5 Pulse Shaping . . . 172
5.6 Symbol Synchronization and Carrier Phase Recovery . . . 175
5.6.1 Receiver Structure with Phase and Timing Recovery . . . 176
5.6.2 Maximum Likelihood Phase Estimation . . . 178
5.6.3 Maximum Likelihood Timing Estimation . . . 179

6 Performance of Digital Modulation over Wireless Channels 187
6.1 AWGN Channels . . . 187
6.1.1 Signal-to-Noise Power Ratio and Bit/Symbol Energy . . . 187
6.1.2 Error Probability for BPSK and QPSK . . . 188
6.1.3 Error Probability for MPSK . . . 190
6.1.4 Error Probability for MPAM and MQAM . . . 192
6.1.5 Error Probability for FSK and CPFSK . . . 194
6.1.6 Error Probability Approximation for Coherent Modulations . . . 195
6.1.7 Error Probability for Differential Modulation . . . 195
6.2 Alternate $\boldsymbol{Q}$-Function Representation . . . 197
6.3 Fading . . . 197
6.3.1 Outage Probability . . . 198
6.3.2 Average Probability of Error . . . 199
6.3.3 Moment Generating Function Approach to Average Error Probability . . . 200
6.3.4 Combined Outage and Average Error Probability . . . 206
6.4 Doppler Spread . . . 207
6.5 Intersymbol Interference . . . 209

7 Diversity 220
7.1 Realization of Independent Fading Paths . . . 220
7.2 Receiver Diversity . . . 222
7.2.1 System Model . . . 222
7.2.2 Selection Combining . . . 224
7.2.3 Threshold Combining . . . 227
7.2.4 Equal-Gain Combining . . . 229
7.2.5 Maximal-Ratio Combining . . . 230
7.2.6 Generalized Combining . . . 232
7.3 Transmitter Diversity . . . 233
7.3.1 Channel Known at Transmitter . . . 233

7.3.2 Channel Unknown at Transmitter – The Alamouti Scheme . . . 234
7.4 Moment Generating Functions in Diversity Analysis . . . 235
7.4.1 Diversity Analysis for MRC . . . 236
7.4.2 Diversity Analysis for EGC, SC, and GC . . . 239
7.4.3 Diversity Analysis for Noncoherent and Differentially Coherent Modulation . . . 240

8 Coding for Wireless Channels 247
8.1 Overview of Code Design . . . 248
8.2 Binary Linear Block Codes . . . 250
8.2.1 Binary Vector Spaces and Hamming Distance . . . 251
8.2.2 Generator Matrix . . . 252
8.2.3 Parity-Check Matrix and Syndrome Testing . . . 255
8.2.4 Cyclic Codes . . . 256
8.2.5 Hard Decision Decoding (HDD) . . . 259
8.2.6 Probability of Error for HDD in AWGN . . . 260
8.2.7 Probability of Error for SDD in AWGN . . . 262
8.2.8 Hamming, Golay, and BCH Codes . . . 264
8.3 Nonbinary Block Codes: The Reed Solomon Code . . . 264
8.4 Convolutional Codes . . . 266
8.4.1 Code Characterization: Trellis Diagrams . . . 266
8.4.2 Maximum Likelihood Decoding . . . 268
8.4.3 The Viterbi Algorithm . . . 271
8.4.4 Distance Properties . . . 272
8.4.5 State Diagrams and Transfer Functions . . . 273
8.4.6 Error Probability for Convolutional Codes . . . 275
8.5 Concatenated Codes . . . 277
8.6 Turbo Codes . . . 278
8.7 Low-Density Parity-Check Codes . . . 281
8.8 Polar Codes . . . 282
8.9 Coded Modulation . . . 285
8.10 Coding with Interleaving for Fading Channels . . . 288
8.10.1 Block Coding with Interleaving . . . 289
8.10.2 Convolutional Coding with Interleaving . . . 291
8.10.3 Coded Modulation with Symbol/Bit Interleaving . . . 291
8.11 Unequal Error Protection Codes . . . 292
8.12 Joint Source and Channel Coding . . . 294

9 Adaptive Modulation and Coding 306
9.1 Adaptive Transmission System . . . 307
9.2 Adaptive Techniques . . . 308
9.2.1 Variable-Rate Techniques . . . 308
9.2.2 Variable-Power Techniques . . . 309
9.2.3 Variable Error Probability . . . 310
9.2.4 Variable-Coding Techniques . . . 310
9.2.5 Hybrid Techniques . . . 310
9.3 Variable-Rate Variable-Power MQAM . . . 310
9.3.1 Error Probability Bounds . . . 311

9.3.2 Adaptive Rate and Power Schemes . . . 312
9.3.3 Channel Inversion with Fixed Rate . . . 313
9.3.4 Discrete-Rate Adaptation . . . 314
9.3.5 Average Fade Region Duration . . . 319
9.3.6 Exact versus Approximate Bit Error Probability . . . 321
9.3.7 Channel Estimation Error and Delay . . . 322
9.3.8 Adaptive Coded Modulation . . . 325
9.4 General M-ary Modulations . . . 326
9.4.1 Continuous-Rate Adaptation . . . 326
9.4.2 Discrete-Rate Adaptation . . . 330
9.4.3 Average BER Target . . . 331
9.5 Adaptive Techniques in Combined Fast and Slow Fading . . . 334

10 MIMO Communications **342**
10.1 Narrowband MIMO Model . . . 342
10.2 Spatial Dimensions of the MIMO Channel . . . 344
10.2.1 Singular Value Decomposition of the Channel Matrix . . . 344
10.2.2 Transmit Precoding and Receiver Combining . . . 344
10.3 MIMO Channel Capacity . . . 346
10.3.1 Static Channels . . . 346
10.3.2 Fading Channels . . . 350
10.4 MIMO Diversity Gain: Beamforming . . . 355
10.5 Diversity–Multiplexing Trade-offs . . . 356
10.6 MIMO Detection Algorithms . . . 357
10.6.1 Maximum Likelihood (ML) Detection . . . 358
10.6.2 Linear Receivers: ZF and MMSE . . . 358
10.6.3 Sphere Decoders (SD) . . . 359
10.7 MIMO Space-Time Codes . . . 360
10.7.1 Pairwise Error Probability . . . 360
10.7.2 Rank and Determinant Design Criteria . . . 361
10.7.3 Space-Time Trellis and Block Codes . . . 362
10.7.4 Spatial Multiplexing and BLAST Architectures . . . 362
10.8 Frequency-Selective MIMO Channels . . . 364

11 Equalization **378**
11.1 ISI Mitigation and Noise Enhancement . . . 379
11.2 Equalizer Types and Structures . . . 379
11.3 Folded Spectrum and ISI-Free Transmission . . . 380
11.4 Linear Equalizers . . . 383
11.4.1 Zero-Forcing (ZF) Equalizers . . . 384
11.4.2 Minimum Mean-Square Error (MMSE) Equalizers . . . 385
11.5 Nonlinear Equalizers . . . 388
11.5.1 Maximum Likelihood Sequence Estimation . . . 388
11.5.2 Decision-Feedback Equalization . . . 390
11.5.3 Turbo and Deep-Learning Equalization . . . 392
11.6 Adaptive Equalizers: Training and Tracking . . . 392

12 Multicarrier Modulation **402**
12.1 Data Transmission Using Multiple Carriers . . . 403
12.2 Multicarrier Modulation with Overlapping Subchannels . . . 405
12.3 Mitigation of Subcarrier Fading . . . 408
12.3.1 Coding with Interleaving over Time and Frequency . . . 408
12.3.2 Frequency Equalization . . . 408
12.3.3 Precoding . . . 409
12.3.4 Pulse Shaping . . . 409
12.3.5 Adaptive Loading . . . 409
12.4 Discrete Implementation of Multicarrier Modulation . . . 410
12.4.1 The DFT and Its Properties . . . 410
12.4.2 The Cyclic Prefix . . . 411
12.4.3 Orthogonal Frequency-Division Multiplexing (OFDM) . . . 413
12.4.4 Matrix Representation of OFDM . . . 415
12.4.5 Vector Coding . . . 417
12.5 Challenges in Multicarrier Systems . . . 420
12.5.1 Peak-to-Average Power Ratio . . . 420
12.5.2 Frequency and Timing Offset . . . 422
12.5.3 Large Doppler Spread . . . 423
12.6 Case Study: OFDM design in the Wi-Fi Standard . . . 424
12.7 Orthogonal Time-Frequency-Space Modulation (OTFS) . . . 425

13 Spread Spectrum **435**
13.1 Spread-Spectrum Principles . . . 435
13.2 Direct-Sequence Spread Spectrum (DSSS) . . . 440
13.2.1 DSSS System Model . . . 440
13.2.2 Spreading Codes for ISI Rejection: Random, Pseudorandom, and m-Sequences . . . 444
13.2.3 Synchronization . . . 447
13.2.4 RAKE Receivers . . . 449
13.3 Frequency-Hopping Spread Spectrum (FHSS) . . . 451
13.4 Multiuser DSSS Systems . . . 453
13.4.1 Spreading Codes for Multiuser DSSS . . . 454
13.4.2 Downlink Channels . . . 457
13.4.3 Uplink Channels . . . 461
13.4.4 Multiuser Detection . . . 466
13.4.5 Multicarrier CDMA . . . 468
13.5 Multiuser FHSS Systems . . . 470

14 Multiuser Systems **480**
14.1 Multiuser Channels: The Uplink and Downlink . . . 480
14.2 Multiple Access . . . 482
14.2.1 Frequency-Division Multiple Access (FDMA) . . . 483
14.2.2 Time-Division Multiple Access (TDMA) . . . 484
14.2.3 Code-Division Multiple Access (CDMA) . . . 485
14.2.4 Space-Division Multiple Access (SDMA) . . . 487
14.2.5 Non-orthogonal Multiple Access (NOMA) . . . 487
14.2.6 Hybrid Techniques . . . 488

14.2.7 Full-Duplex Systems . . . 489
14.3 Random Access . . . 491
14.3.1 Pure ALOHA . . . 492
14.3.2 Slotted ALOHA . . . 493
14.3.3 Carrier-Sense Multiple Access (CSMA) . . . 494
14.4 Scheduling . . . 495
14.5 Power Control . . . 496
14.6 Downlink (Broadcast) Channel Capacity . . . 498
14.6.1 Channel Model . . . 499
14.6.2 Capacity in AWGN . . . 500
14.6.3 Common Data . . . 505
14.6.4 Capacity in Fading . . . 507
14.6.5 Capacity with Multiple Antennas . . . 511
14.7 Uplink (Multiple Access) Channel Capacity . . . 513
14.7.1 Capacity in AWGN . . . 513
14.7.2 Capacity in Fading . . . 516
14.7.3 Capacity with Multiple Antennas . . . 517
14.8 Uplink-Downlink Duality . . . 518
14.9 Multiuser Diversity . . . 521
14.10 Multiuser MIMO Systems . . . 523

15 Cellular Networks **535**
15.1 Design Principles . . . 536
15.1.1 Channel Reuse and SINR . . . 536
15.1.2 Interference Reduction Techniques . . . 537
15.1.3 Cell Shape and Size . . . 538
15.1.4 Channel Allocation and Reuse Distance . . . 539
15.1.5 Base Station Antenna Design . . . 542
15.2 SIR and User Capacity . . . 544
15.2.1 Orthogonal Systems (TDMA/FDMA) . . . 545
15.2.2 Nonorthogonal Systems (CDMA) . . . 547
15.3 Dynamic Resource Allocation . . . 549
15.3.1 Dynamic Channel Assignment . . . 549
15.3.2 Power Control . . . 550
15.4 Advanced Network Architectures . . . 551
15.4.1 Heterogeneous Networks (Hetnets) . . . 551
15.4.2 Cloud Radio Access Networks (C-RAN) . . . 553
15.5 Fundamental Rate Limits . . . 554
15.5.1 Shannon Capacity of Cellular Systems . . . 554
15.5.2 Area Spectral Efficiency . . . 555

16 Ad Hoc Wireless Networks **568**
16.1 Design Principles . . . 569
16.1.1 Protocol Layers . . . 570
16.1.2 Physical Layer Design: Point to Point Communications . . . 572
16.1.3 Access Layer Design: Multiple Access . . . 572
16.1.4 Network Layer Design: Discovery, Routing, and Resource Allocation . . . 574

16.1.5 Transport Layer Design: End-to-End Communication . . . 578
16.1.6 Application Layer Design . . . 578
16.1.7 Cross-Layer Design . . . 579
16.2 Routing Protocols . . . 581
16.3 Capacity Limits . . . 584
16.4 Energy-Constrained Networks . . . 585
16.4.1 Modulation and Coding . . . 586
16.4.2 MIMO and Cooperative MIMO . . . 587
16.4.3 Access, Routing, and Sleeping . . . 588
16.4.4 Cross-Layer Design under Energy Constraints . . . 589
16.4.5 Capacity per Unit Energy . . . 589
16.5 Network Types . . . 590
16.5.1 Data Networks . . . 591
16.5.2 Sensor and Distributed Learning Networks . . . 591
16.5.3 Mobile Ad Hoc Networks . . . 592
16.5.4 Distributed Control Systems . . . 593

A Representation of Bandpass Signals and Channels 605

B Probability Theory, Random Variables, and Random Processes 609
B.1 Probability Theory . . . 609
B.2 Random Variables . . . 610
B.3 Random Processes . . . 614
B.4 Gaussian Processes . . . 617

C Matrix Definitions, Operations, and Properties 620
C.1 Matrices and Vectors . . . 620
C.2 Matrix and Vector Operations . . . 621
C.3 Matrix Decompositions . . . 624

D Summary of Wireless Standards 627
D.1 Cellular Phone Standards . . . 627
D.2 Wi-Fi Standards . . . 634
D.3 Short-Range Wireless Network Standards . . . 638

Chapter 1

Overview of Wireless Communications

Wireless communication is one of the most impactful technologies in history, drastically affecting the way we live, work, play, and interact with people and the world. There are billions of cellphone subscribers worldwide, and a wide range of devices in addition to phones use cellular technology for their connectivity. Wireless network technology using the Wi-Fi standard has been incorporated into billions of devices as well, including smartphones, computers, cars, drones, kitchen appliances, watches, and tennis shoes. Satellite communication systems support video, voice, and data applications for receivers on earth, in the air, and in space. Revenue across all areas of wireless technology and services is trillions of dollars annually. The insatiable demand for wireless data along with new and compelling wireless applications indicate a bright future for wireless systems. However, many technical challenges remain in designing wireless networks and devices that deliver the performance necessary to support existing and emerging applications. In this introductory chapter we will briefly review the history of wireless communications, from the smoke signals of antiquity to the rise of radio communication that underlies the Wi-Fi, cellular, and satellite networks of today. We then discuss the most prevalent wireless communication systems in operation today. The impact of spectrum properties and regulation as well as standards on the design and success of wireless systems is also illuminated. We close the chapter by presenting a vision for the wireless communication systems of the future, including the technical challenges that must be overcome to make this vision a reality. Techniques to address many of these challenges are covered in subsequent chapters. The huge gap between the capabilities of current systems and the vision for future systems indicates that much research and development in wireless communications remain to be done.

1.1 History of Wireless Communications

This section describes the early history of radio communications, its evolution from analog to digital technology, and its incorporation into current wireless systems and standards.

1.1.1 Origins of Radio Technology

The first wireless networks were developed in antiquity. These systems transmitted information visually over line-of-sight distances (later extended by telescopes) using smoke signals, torch signaling, flashing mirrors, signal flares, or semaphore flags. An elaborate set of signal combinations was developed to convey complex messages with these rudimentary signals. Observation stations were built on hilltops and along roads to relay these messages over large distances. These early communication networks were replaced first by the telegraph network (invented by Samuel Morse in 1838) and later by the telephone.

The origins of radio communications began around 1820 with experiments by Oersted demonstrating that an electric field could move a compass needle, thereby establishing a connection between electricity and magnetism. Work by Ampére, Gauss, Henry, Faraday, and others further advanced knowledge about electromagnetic waves, culminating in Maxwell's theory of electromagnetism published in 1865. The first transmission of electromagnetic waves was performed by Hertz in the late 1880s, after which he famously declared to his students that these waves would be "of no use whatsoever." He was proved wrong in 1895 when Marconi demonstrated the first radio transmission across his father's estate in Bologna. That transmission is considered the birth of radio communications, a term coined in the early 1900s. Marconi moved to England to continue his experiments over increasingly large transmission ranges, culminating in the first trans-Atlantic radio transmission in 1901. In 1900 Fessenden became the first person to send a speech signal over radio waves, and six years later he made the first public radio broadcast. From these early beginnings, described in more detail in [1], radio technology advanced rapidly to enable transmissions over larger distances with better quality, less power, and smaller, cheaper devices, thereby enabling public and private radio communications, television, and wireless networking.

1.1.2 From Analog to Digital

Radio systems designed prior to the invention of the transistor, including AM/FM radio, analog television, and amateur radio, transmitted analog signals. Most modern radio systems transmit digital signals generated by digital modulation of a bit stream. The bit stream may represent binary data (e.g., a computer file, digital photo, or digital video stream) or it may be obtained by digitizing an analog signal (e.g., by sampling the analog signal and then quantizing each sample). A digital radio can transmit a continuous bit stream or it can group the bits into packets. The latter type of radio is called a *packet radio* and is characterized by bursty transmissions: the radio is idle except when it transmits a packet. When packet radios transmit continuous data such as voice and video, the delay between received packets must not exceed the delay constraint of the data. The first wireless network based on packet radio, ALOHAnet, was developed at the University of Hawaii and began operation in 1971. This network enabled computer sites at seven campuses spread out over four islands to communicate with a central computer on Oahu via radio transmission. The network architecture used a star topology with the central computer at its hub. Any two computers could establish a bi-directional communications link between them by going through the central hub. ALOHAnet incorporated the first set of protocols for channel access and routing in packet radio systems, and many of the underlying principles in these protocols are still in use today.

The U.S. military saw great potential for communication systems exploiting the combination of packet data and broadcast radio inherent to ALOHAnet. Throughout the 70's and early 80's the Defense Advanced Research Projects Agency (DARPA) invested significant resources to develop networks using packet radios for communications in the battlefield. The nodes in these packet radio networks had the ability to configure (or reconfigure) into a network without the aid of any established infrastructure. Self-configuring wireless networks without any infrastructure were later coined *ad hoc* wireless networks.

DARPA's investment in packet radio networks peaked in the mid 1980's, but these networks fell far short of expectations in terms of speed and performance. This was due in part to the limited capabilities of the radios and in part to the lack of robust and efficient access and routing protocols. Packet radio networks also found commercial application in supporting wide-area wireless data services. These services, first introduced in the early 1990's, enabled wireless data access (including email, file transfer, and web browsing) at fairly low speeds, on the order of 20 Kbps. The market for these wide-area wireless data services did not take off due mainly to their low data rates, high cost, and lack of "killer applications". All of these services eventually folded, spurred in part by the introduction of wireless data in 2G cellular services [2], which marked the dawn of the wireless data revolution.

1.1.3 Evolution of Wireless Systems and Standards

The ubiquity of wireless communications has been enabled by the growth and success of Wireless Local Area Networks (WLANs), standardized through the family of IEEE 802.11 (Wi-Fi) protocols, as well as cellular networks. Satellite systems also play an important role in the wireless ecosystem. The evolution of these systems and their corresponding standards is traced out in this subsection.

Wi-Fi Systems

The success story of Wi-Fi systems, as illuminated in [3], began as an evolution of the Ethernet (802.3) standard for wired local area networks (LANs). Ethernet technology, developed at Xerox Parc in the 1970s and standardized in 1983, was widely adopted throughout the 1980s to connect computers, servers, and printers within office buildings. WLANs were envisioned as Ethernet LANs with cables replaced by radio links. In 1985 the Federal Communications Commission (FCC) enabled the commercial development of WLANs by authorizing for unlicensed use three of the Industrial, Scientific, and Medical (ISM) frequency bands: the 900 MHz band spanning 902-928 MHz, the 2.4 GHz band spanning 2.4-2.4835 GHz, and the 5.8 GHz band spanning 5.725-5.875 GHz. Up until then, these frequency bands had been reserved internationally for radio equipment associated with industrial, scientific and medical purposes other than telecommunications. Similar rulings followed shortly thereafter from the spectrum regulatory bodies in other countries. The new rulings allowed unlicensed use of these ISM bands by any radio following a certain set of restrictions to avoid compromising the performance of the primary band users. Such radios were also subject to interference from these primary users. The opening of these ISM bands to "free" use by unlicensed wireless devices unleashed a flurry of research and commercial wireless system development, particularly for WLANs.

The first WLAN product for the ISM band, called WaveLAN, was launched in 1988 by the NCR corporation and cost several thousand dollars. These WLANs had data rates up to 2 Mbps and operated in the 900 MHz and 2.4 GHz ISM bands. Dozens of WLAN companies and products appeared over the ensuing few years, mostly operating in the 900 MHz ISM band using direct-sequence spread spectrum, with data rates on the order of 1-2 Mbps. The lack of standardization for these products led to high development costs, poor reliability, and lack of interoperability between systems. Moreover, Ethernet's 10 Mbps data rate and high reliability far exceeded the capabilities of these early WLAN products. Since companies were willing to run cables within and between their facilities to get this better performance, the WLAN market remained small for its first decade of existence. The initial WLAN products were phased out as the 802.11 standards-based WLAN products hit the market in the late 1990s. In Europe a WLAN standard called HIPERLAN was finalized in 1996. An improved version was launched in 2000, but HIPERLAN systems never gained much traction.

The first 802.11 standard, finalized in 1997, was born from a working group within the IEEE 802 LAN/MAN Standards Committee that was formed in 1990. This standard, called 802.11-1997, operated in the 2.4 GHz ISM band with data rates up to 2 Mbps. The standard used channel sensing with collision avoidance for medium access as well as frequency hopping or direct-sequence spread spectrum to mitigate the main sources of interference in that band which, at the time, consisted primarily of microwave ovens, cordless phones, and baby monitors. Also in 1997, the FCC authorized 200 MHz of additional unlicensed spectrum in the 5 GHz band, with certain constraints to avoid interfering with the primary users, mostly radar systems, operating in this band. Two years later, in 1999, the standard that would ignite the Wi-Fi revolution, 802.11b, was finalized. This standard increased data rates to 11 Mbps and eliminated the frequency-hop option of its predecessor. A plethora of 802.11b products soon appeared. Their interoperability, coupled with dramatically lower costs relative to earlier WLAN products, led to widespread use of this new technology. Millions of 802.11b products were shipped in 2000, just one year after the standard was finalized, and these shipments grew tenfold by 2003. The year 1999 marked two other important milestones for WLANs. That year the 802.11a standard for the 5GHz ISM frequency band was finalized to capitalize on the new

unlicensed spectrum in that band. This standard enabled 54 Mbps data rates in 20 MHz channels and introduced orthogonal-frequency-division-multiplexing (OFDM) coupled with adaptive modulation as a new physical layer design. Also in 1999 the Wireless Ethernet Compatibility Alliance was formed to facilitate interoperability and certification of WLAN products. The name of this group was later abbreviated to the Wi-Fi Alliance, thereby coining the widely used moniker for WLAN technology and standards today.

The Wi-Fi standard has evolved rapidly since its early days, with new versions developed about every five years, as described in more detail in Appendix D.1. The 802.11g standard, introduced in 2003, is essentially the same in its physical layer and multiple access design as 802.11a, but operates in the 2.4 GHz ISM frequency band. Later standards provided improvements over these first OFDM Wi-Fi systems, including wider channels, multiple transmit and receive antennas to enable multiple spatial streams and improved robustness through beamforming, larger signal constellations, improved error-correction codes, and coordinated multiple access. These improvements have led to highly-reliable Wi-Fi products that are capable of 10 Gbps data rates within a building or outdoor area. Wi-Fi has also moved into the unregulated 60 GHz frequency band through the 802.11ad standard. Even its name has evolved; In 2018 the 802.11 standards body abandoned using letter suffixes for new generations of the standard. Instead it coined the sixth generation of the WiFi standard, originally named 802.11ax, as Wi-Fi 6. Today Wi-Fi technology is pervasive indoors and out even in remote corners of the world. In addition to its pervasiveness, Wi-Fi has experienced an explosion of applications beyond its original use of connecting computers to each other and their peripheral devices. In addition to computers, smartphones, and tablets, many electronic devices today, from medical devices to refrigerators to cars, are equipped with Wi-Fi, allowing them to download new software, exchange data with other devices, and take advantage of cloud-based storage and computation.

Cellular Networks

Cellular networks are another exceedingly successful wireless technology. The convergence of radio and telephony began in 1915, when wireless voice transmission between New York and San Francisco was first established. The first analog mobile telephone system was deployed in St. Louis Missouri in 1946, launching AT&T's Mobile Telephone Service (MTS). Within two years AT&T had deployed MTS over approximately 100 cities and highway corridors. Only six channels were allocated by the FCC for the service and, due to their close spacing in frequency, only three were usable at any given time. Hence only three people within a city could make a call simultaneously. The monthly service and per-call cost was very high, and the equipment bulky and heavy. Evolution of the system was slow; while the equipment improved and spectrum to support up to 12 channels was added, the system capacity remained extremely limited.

Ironically, about the same time MTS was first being deployed, a solution to this capacity problem had already emerged from researchers at AT&T Bell Laboratories: the notion of cellular networks. The cellular network concept, articulated in a 1947 Bell Laboratories Technical Memo by D. H. Ring [4], exploited the fact that the power of a transmitted signal falls off with distance. Thus, channels using the same frequency can be allocated to users at spatially-separate locations with minimal interference between the users. To exploit this principle of frequency reuse, a cellular network partitions a geographical area into non-overlapping *cells*, as shown in Fig. 1.3 below. Sets of channels are assigned to each cell, and cells that are assigned the same channel set are spaced far enough apart so that interference between the users in these cells is small. In early cellular networks the distance between cells using the same channel set was relatively large, but today sophisticated interference mitigation techniques allow channels to be reused in every cell. As a user moves between adjacent cells, its call is *handed off* to a channel associated with the new cell. Frequency reuse enables much more efficient use of spectrum as the number of simultaneous users is no longer limited to the number of available channels. Indeed, while many aspects of cellular network technology have changed over time, frequency reuse remains at the heart of cellular network design.

Although the cellular concept was introduced in the late 1940s, it was not implemented for several decades, as technology was not yet ripe to realize the system in practice. In the mid-1960s engineers at Bell Laboratories began work on a design and feasibility study for a metropolitan analog cellular network. The details of the design and analysis, along with successful system tests in Newark and Philadelphia, formed the basis of an AT&T FCC proposal in 1971 to approve cellular service and allocate spectrum for it [5]. The FCC approved experimental cellular licenses to telephone companies in 1974, which led to the construction of several cellular networks. Following a long and convoluted process, in 1981 the FCC finalized its ruling for the issuance of commercial cellular licenses, instituting a duopoly in every metropolitan area whereby one license would be granted to a traditional phone company, and the other to a non-traditional operator. The first generation cellular network for U.S. deployment, called the advanced mobile phone system (AMPS) and described in [6], was launched in Chicago in 1983. A similar service had been launched in Tokyo in 1979 and in Scandinavia in 1981, although those systems were incompatible with AMPS. Cellular service in many other countries launched in the ensuing years, mostly with incompatible standards, which precluded cellular roaming between countries. These first generation systems are referred to as 1G systems.

Like Wi-Fi, the exponential growth of the cellular industry exceeded all expectations, increasing by at least an order of magnitude every decade, from a modest 100,000 subscribers in 1984 to tens of millions in 1994, hundreds of millions in 2004, and billions in 2014. To meet the growing demand for wireless data along with a diverse requirements for different types of wireless devices, a new generation of cellular network standards has emerged approximately every decade. In fact, that timeline was compressed for 1G systems, since the analog cellular network deployed in Chicago in 1983 was already saturated by 1984. At that point the FCC increased the cellular spectral allocation from 40 MHz to 50 MHz. As cellular networks throughout more and more cities became saturated with demand, the development of digital cellular technology for increased capacity and better performance became essential. These enhanced system requirements launched the process to create a second generation (2G) cellular standard in the late 1980s with deployments in the early 1990s. In addition to voice communication, the move to digital technology paved the way for these systems to support low-rate data, on the order of 20 Kbps for upload and 100 kbps for download. The demand for digital cellular technology worldwide led to a proliferation of 2G cellular standards, with three different standards in the U.S. alone. One of these matched the European standard, which China also adopted. Japan adopted a separate standard. Hence global roaming required a multi-mode phone.

The third generation of cellular standards (3G) aspired to create a single worldwide standard to drive economies of scale, technology innovation, and seamless roaming. To do so, seven telecommunications standards bodies across Asia, Europe and the United States formed the third-generation partnership project (3GPP). The 3G cellular networks based on the 3GPP standard, whose deployments began in the early 2000s, provided an order of magnitude higher data rates than 2G systems, on the order of 1 Mbps for upload and 3-7 Mbps for download. Indeed, it was the capabilities of 3G systems that transitioned the "killer application" of cell phones from voice to wireless data. The proliferation of smart phone technology in the mid-2000s, which were designed to consume vast amounts of wireless data, greatly stressed the capacity of 3G networks. Moreover, the 3G networks had far lower data rates than the 54 Mbps offered by the Wi-Fi standard at the time. Latency was also high, up to 500 ms, which precluded interactive applications like gaming and video calls. Hence, to enable high-speed data and real-time interaction, users typically switched their phones from cellular to Wi-Fi whenever possible. These developments paved the way for a rush to develop the next generation cellular standard to provide similar performance as Wi-Fi.

The 4G cellular standard set data rate requirements at 100 Mbps for mobile users and 1 Gbps for stationary users, with latencies below 100 ms to enable interactive applications. As 4G standardization to meet these requirements was progressing in the late 2000s, carriers upgraded their networks to the "long term evolution (LTE)" standard which provided better performance than 3G but did not meet minimum requirements of 4G. Simultaneously, an alternative standard to LTE called WiMAX was developed to meet 4G requirements, but it never gained

sufficient traction to compete with LTE [5]. Deployment of 4G networks based on the LTE standard became widespread by the mid-2010s. In addition to an order of magnitude data rate improvement, these networks reduced latency to around 20 ms by introducing packet-switched networking versus the voice-oriented circuit-switched architecture of prior cellular networks. In the late 2010s the 4G standard was enhanced with multiple antenna (MIMO) technology and carrier aggregation to increase data rates 5-10 fold; these enhancements were coined LTE Advanced or 4G+. The high data rates and low latency of 4G and 4G+ networks meet the requirements for many smart phone applications, including video streaming, video calls, and gaming. However, higher speeds and lower latency are always desirable to support more users, faster data downloads, and more demanding applications. These considerations, along with the promise of lower operating costs to carriers, led to 5G standardization. The 5G standard, first deployed in the early 2020s, was the first cellular standard to utilize millimeter wave spectrum in order to support data rates up to 10 Gbps with latencies below 1 ms. In the lower frequency bands data rates and latency are similar to those of 4G+. The 6G cellular standard promises another order of magnitude increase in data rates and decrease in latency. That would continue the trend of data rates increasing by an order of magnitude with each new generation of the cellular standard that started with the evolution from 2G to 3G.

Starting with the 3G systems, each generation of cellular standard has been developed to meet a set of International Mobile Telecommunications (IMT) requirements outlined by the International Telecommunications Union (ITU). The ITU certifies which cellular standards meet its IMT requirements; the 3GPP standard is designed to meet these requirements, as are alternate standards developed by specific countries or other standards organizations. More details on current cellular technology will be given in Chapter 15, with the evolution of cellular standards described in Appendix D.2.

Satellite Systems

Satellite communication systems are another major component of today's wireless communications infrastructure. Commercial satellite systems can provide broadcast services over very wide areas, fill the coverage gap in locations without cellular service, and provide connectivity for aerial systems such as airplane Wi-Fi. Satellite systems are typically categorized by the height of the satellites' orbit: low-earth orbit (LEO) satellites operate at roughly 2000 km altitude, medium-earth orbit (MEO) satellites at roughly 9000 km altitude, and geosynchronous orbit (GEO) satellites at roughly 40,000 km altitude. GEO satellites with geostationary orbits are seen as stationary from the earth, whereas GEO satellites with other orbits (such as elliptical) have their coverage area change over time. The disadvantage of high altitude orbits is that it takes a great deal of power to reach the satellite, and the propagation delay is typically too large for two-way delay-constrained applications like voice, video, and gaming. However, satellites at these orbits tend to have larger coverage areas, so fewer satellites are necessary to provide wide-area or global coverage. In addition to communication services, satellites are used for many other applications including weather monitoring, earth observation, surveillance, imaging, navigation, and localization.

The concept of using GEO satellites for communications was first suggested by Herman Potocnik in 1928 and later popularized through a 1945 article in *Wireless World* written by the science fiction writer Arthur C. Clarke. However, the first deployed satellites, the Soviet Union's Sputnik in 1957 and the Nasa/Bell Laboratories' Echo-1 in 1960, were not geosynchronous due to the difficulty of lifting a satellite into such a high orbit. Following these launches, in 1962 the Communications Satellite Corporation (Comsat) was formed in the United States to develop commercial communication satellite systems. Two years later, the International Telecommunications Satellite Consortium (IntelSat) emerged as a public-private consortium of 18 countries with the goal of enabling global telecommunications connectivity. SYNCOM 2, the first communication satellite to successfully reach geosynchronous orbit, was launched in 1963. The following year SYNCOM 3 was launched into geostationary orbit, providing a two-way 10 MHz communication channel at a carrier frequency of 1.8 GHz for the satellite-to-earth link (downlink) and at 7.3 GHz for the reverse link (uplink). Shortly after its launch, SYNCOM 3 was used to

provide live television coverage to US viewers of the 1964 Summer Olympics in Tokyo. The IntelSat consortium launched a number of satellites during the late 1960s in order to reach near-global coverage, culminating in the broadcasting of the 1969 landing of the first human on the moon to 600 million viewers. In addition to television broadcasting, these early systems also supported voice, teletype, and paging.

GEO satellite technology has evolved continuously since its early days. The C band of the radio spectrum (4-8 GHz) was used in early GEO systems, but subsequent systems have taken advantage of the larger bandwidths available at the higher frequency X (8-12 GHz), Ku (12-18 GHz), K (18-26 GHz) and Ka (26-40 GHz) bands in order to provide higher data rates. Antenna system technology for GEOs has also evolved to focus the transmission energy more precisely, leading to higher signal quality and hence higher data rates. There are hundreds of GEO satellites deployed today, with some capable of a total data rate in excess of one hundred Gbps. For individual users, typical data rates are on the order of ten Mbps in the downlink and several Mbps in the uplink [7]. Due to their large coverage regions, geosynchronous satellites are the primary communications mechanism for vessels at sea and in the air. They are also well-suited for broadcast entertainment where round-trip delay is not a consideration; they support hundreds of channels of digital television and radio at very high quality.

As cellular networks were rolled out in the late 1980s, the focus for satellite technology turned to building LEO systems that might compete with them. This led to the launch of several LEO communication systems in the late 1990s, including Globalstar and Iridium [8]. These LEOs provided global coverage but the link rates remained low due to power and bandwidth constraints. The handsets for these systems were also much larger and heavier than their cellular counterparts, primarily due to their large batteries. The LEO satellite systems deployed in the 1990s did not experience significant commercial success as, for most users, they provided worse performance and coverage than the competing 2G cellular phone technology at a higher cost. While these satellite systems provided coverage to remote areas without access to cellular service, the cost was prohibitive for the large majority of people in these areas. As a result, pretty much all the initial LEO companies, including Globalstar and Iridium, filed for bankruptcy protection within a few years of launching their service. Teledesic, which was planning a very ambitious LEO system with more than 800 satellites, declared bankruptcy before a single launch. After several iterations of restructuring, Iridium and Globalstar emerged as profitable services and today provide global Internet connectivity, albeit for a relatively small number of subscribers compared to that of cellular service. Interest in LEO satellites resurfaced as 4G cellular networks rolled out due to increased demand for connectivity in remote locations not served by cellular, better technology for both satellites and ground transceivers, and reduced launch costs. Significant commercial development of LEO satellites emerged concurrent with 5G cellular rollouts as a means to provide broadband Internet service to the billions of people worldwide that do not have access to high-speed connectivity. It is anticipated that such systems will each deploy hundreds to thousands of satellites.

Today we take for granted the magic of wireless communications, which allows us and our devices to be connected anywhere in the world. This magic has been enabled by the pioneers that contributed to advances in wireless technology underlying the powerful systems ubiquitous today.

1.2 Wireless Network Paradigms

This section provides an overview of the most prevalent wireless systems in operation today. The design details of these systems are constantly evolving, incorporating new technologies and innovations. This section will focus mainly on the high-level design aspects of these systems. More details on these systems and their underlying technologies will be provided in specific sections of the book. A summary of Wi-Fi, cellular, and short-range networking standards along with references for more details can be found in Appendix D.

1.2.1 Wireless Local Area Networks

WLANs support data transmissions for multiple users within a "local" coverage area whose size depends on the radio design, operating frequency, propagation characteristics, and antenna capabilities. A WLAN consists of one or more access points (APs) connected to the Internet that serve one or more WLAN devices, also called clients. The basic WLAN architecture is shown in Figure 1.1. In this architecture clients within the coverage area of an AP connect to it via a single-hop radio transmission.

A WLAN may incorporate multiple APs connected together over wireless links as well as dedicated relay nodes. These enhanced architectural features enable multi-hop transmissions between clients and APs, as shown in Figure 1.2. Relaying data through an intermediate node generally results in better coverage than direct transmission, at a cost of worse spectral efficiency and larger delay. This performance degradation is due to the processing delay associated with relaying and the *half-duplex* nature of most relay nodes, which means they cannot transmit and receive on the same channel at the same time. Hence, each half-duplex relay in a multihop route reduces the data rate by half. A full-duplex relay simultaneously transmits and receives on the same channel at the same time and hence has no spectral efficiency penalty relative to direct transmission. Yet full duplex relaying entails other performance and cost challenges that have precluded their widespread use, as described in Chapter 14.2.7.

WLANs almost exclusively follow the 802.11 Wi-Fi family of protocols and hence are referred to as *Wi-Fi networks*. A Wi-Fi network with multiple APs connected together wirelessly is referred to as a *Wi-Fi mesh network*. Handoff of a moving client between APs is not supported by the 802.11 protocol, however some WLAN networks, particularly corporate and campus systems, incorporate this feature into their overall design.

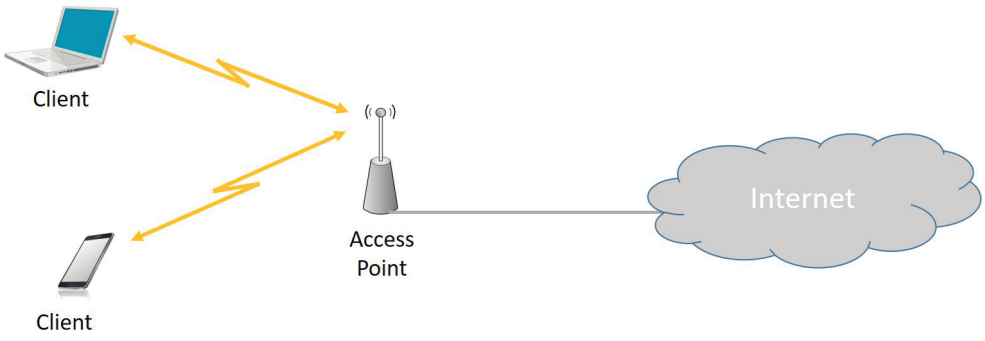


Figure 1.1: Basic wireless LAN architecture

While early WLANs were exclusively indoor systems, current WLANs operate both indoors and outdoors. Indoor WLANs are prevalent in most homes, offices, and other indoor locations where people congregate. Outdoor systems are generally deployed in areas with a high density of users such as corporate and academic campuses, sports stadiums, and downtown areas. The range of an indoor WLAN is typically less than 50 meters and can be confined to a single room for those operating at 60 GHz. Outdoor systems have a bigger range than indoor systems due to their higher power, better antennas, and lower density of obstructions between the transmitter and receiver. WLANs generally operate in unlicensed frequency bands, hence share these bands with other unlicensed devices, such as cordless phones, baby monitors, security systems, and Bluetooth radios. Interference between WLAN devices is controlled by the WLAN access protocol, whereas interference between WLAN devices and other unlicensed devices is mitigated by a limit on the power per unit bandwidth for such devices.

WLANs use packet data transmission for better sharing of the network resources. Hence, data files are segmented into packets. Sharing of the available bandwidth between different APs and, for most systems, between clients accessing the same AP is typically controlled in a distributed manner using the carrier sense multiple access

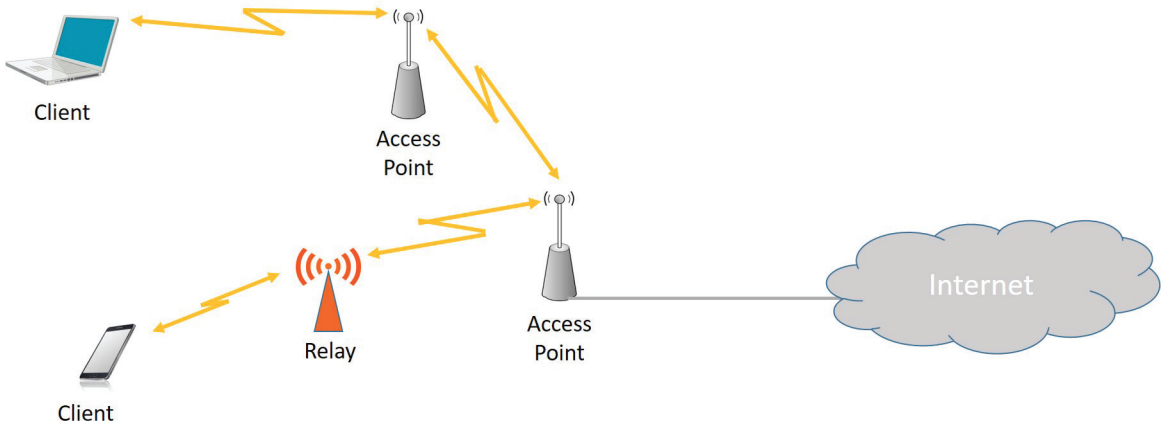


Figure 1.2: Wireless LAN architecture with multihop transmissions

with collision avoidance (CSMA/CA) protocol. In CSMA/CA, energy from other transmissions on a given channel is sensed prior to a transmission and, if the energy is sensed above a given threshold, the transmitter waits a random *backoff* time before again attempting a transmission. A transmitted packet is acknowledged by the receiver once received. If a transmitted packet is not acknowledged, it is assumed lost and hence retransmitted. As discussed in more detail in Chapter 14, the CSMA/CA access protocol is very inefficient, which leads to poor performance of WLANs under moderate to heavy traffic loads. As a result, some WLANs use more sophisticated access protocols, similar to those in cellular networks, that provide centralized scheduling and resource allocation either to all clients served by a given AP or, more generally, to all APs and clients within a given system.

1.2.2 Cellular Networks

Cellular networks provide connectivity to a wide range of devices, both indoors and out, with a plethora of features and applications including text messages, voice calls, data and video transfer, Internet access, and mobile "apps" (application software tailored to run on mobile devices). cellular networks typically operate in licensed frequency bands, whereby the cellular operator must purchase or lease the spectrum in which their system operates. These licenses typically grant the operator exclusive use of the licensed spectrum. Starting in 2017 regulators in several countries began allowing cellular networks to operate in the 5 GHz unlicensed band in addition to within their licensed spectrum. Given the large amount of spectrum in these unlicensed bands compared to what is available in the licensed bands, this development has the potential to drastically increase cellular network capacity. However, it will also increase interference in the unlicensed bands, particularly for Wi-Fi systems.

The basic premise behind cellular network design is frequency reuse, which exploits the fact that signal power falls off with distance to reuse the same frequency spectrum at spatially separated locations. Specifically, a cellular network consists of multiple cells, where each cell is assigned one or more channels to serve its users. Each channel may be reused in another cell some distance away, and the interference between cells operating on the same channel is called *intercell interference*. The spatial separation of cells that reuse the same channel set, the *reuse distance*, should be as small as possible so that frequencies are reused as often as possible, thereby maximizing spectral efficiency. However, as the reuse distance decreases, the intercell interference increases owing to the smaller propagation distance between interfering cells. Early cellular networks had reuse distances greater than one, but current systems typically reuse each channel in every cell while managing the resulting interference through dynamic channel allocation coupled with sophisticated interference mitigation techniques.

The most basic cellular network architecture consists of tessellating cells, as shown in Figure 1.3, where C_i

denotes the channel set assigned to cell i. The two-dimensional cell shapes that tessellate (cover a region without gaps or overlap) are hexagons, squares, and triangles. Of these, the hexagon best approximates an idealized omni-directional transmitter's circular coverage area. Early cellular networks followed this basic architecture using a relatively small number of cells to cover an entire city or region. The cell base stations in these systems were placed on tall buildings, towers, or mountains and transmitted at very high power with cell coverage areas of one or more square kilometers. These large cells are called *macrocells*. The first few generations of macrocell base stations used single-antenna omni-directional transmitters, so a mobile moving in a circle around these base stations had approximately constant average received power unless the signal was blocked by an attenuating object. In deployed systems cell coverage areas overlap or have gaps as signal propagation from a set of base stations never creates tessellating shapes in practice, even if they can be optimally placed. Optimal base station placement is also impractical, as zoning restrictions, rooftop availability, site cost, backhaul and power availability, as well as other considerations influence this placement.

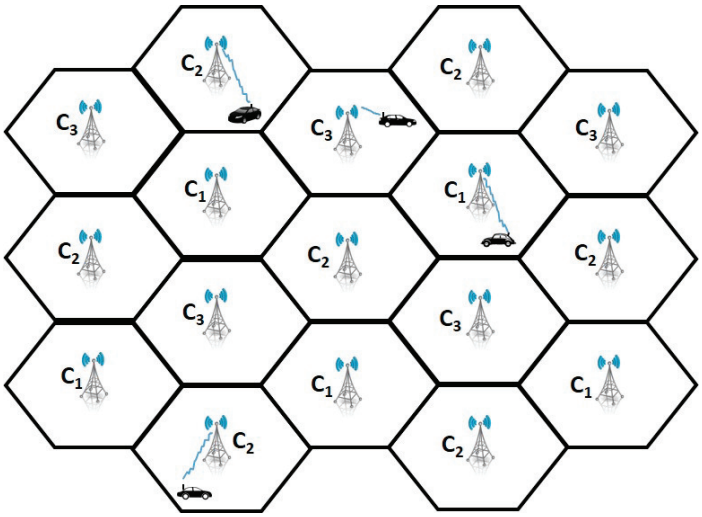


Figure 1.3: Cellular network architecture (homogeneous cell size).

Macrocells have the benefit of wide area coverage, but they can often become overloaded when they contain more users than channels. This phenomenon has led to hierarchical cellular network architectures, with macrocells that provide wide area coverage and small cells embedded within these larger cells to provide high capacity, as shown in Figure 1.4. Not only do the small cells provide increased capacity over a macrocell-only network, but they also reduce the transmit power required at both the base station and mobile terminal, since the maximum transmission distance within the small cell is much less than in the macrocell. Cellular networks with heterogeneous cells sizes are referred to as *Hetnets*. In current cellular networks, channels are typically assigned dynamically based on interference conditions. Another feature of many current cellular network designs is for base stations in adjacent macrocells to operate on the same frequency, utilizing power control, adaptive modulation and coding, as well as interference mitigation techniques to ensure the interference between users does not preclude acceptable performance. Cellular networks can also employ multihop routing between a base station and a mobile terminal by deploying dedicated relays or by enabling intermediate mobile terminals to serve as relays. Multihop routing in cellular networks can provide connectivity for mobiles that have no connectivity to any base station but, as with wireless LANs, generally results in worse performance than single-hop routing due to the delay and data rate reduction incurred by the relays.

All base stations in a given geographical area, both macrocells and small cells, are connected via a high-speed communications link to a mobile telephone switching office (MTSO), as shown in Figure 1.5. The MTSO acts as

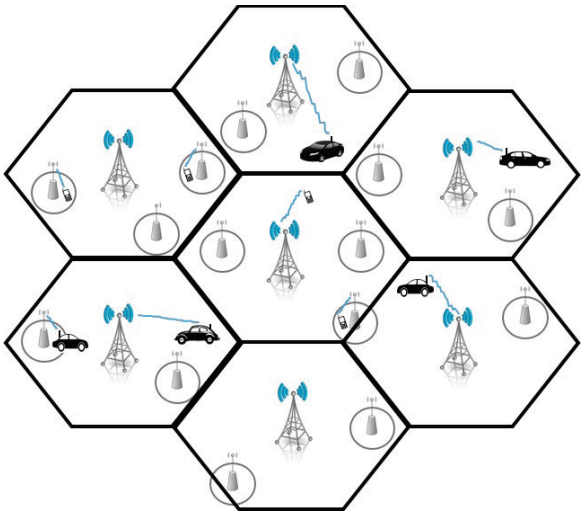


Figure 1.4: Hierarchical cellular network architecture (a Hetnet with macrocells and small cells).

a central controller for the cellular network, allocating channels within each cell, coordinating handoffs between cells when a mobile traverses a cell boundary, and routing calls to and from mobile users. The MTSO can route calls to mobile users in other geographic regions via the local MTSO in that region or to landline users through the public switched telephone network (PSTN). In addition, the MTSO can provide connectivity to the Internet.

A new user located in a given cell requests a channel by sending a call request to the cell's base station over a separate control channel. The request is relayed to the MTSO, which accepts the call request if a channel is available in that cell. If no channels are available then the call request is rejected. A call handoff is initiated when the base station or the mobile in a given cell detects that the received signal power for that call is approaching a given minimum threshold. In this case the base station informs the MTSO that the mobile requires a handoff, and the MTSO then queries surrounding base stations to determine if one of these stations can detect that mobile's signal. If so then the MTSO coordinates a handoff between the original base station and the new base station. If no channels are available in the cell with the new base station then the handoff fails and the call is dropped. A call will also be dropped if the signal strength between a mobile and its base station falls below the minimum threshold needed for communication due to signal propagation effects such as path loss, blockage, or multipath fading. These propagation characteristics are described in Chapters 2-3.

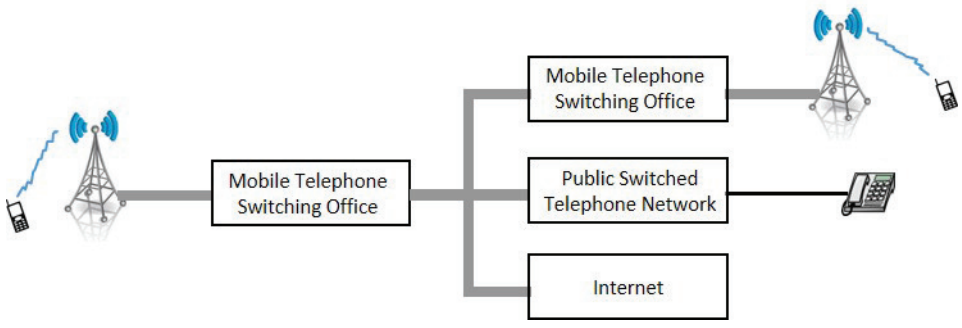


Figure 1.5: Cellular network architecture evolution.

Spectral sharing in cellular networks, also called multiple access, is done by dividing the signaling dimensions along the time, frequency, code, and/or the spatial dimensions. Current cellular standards use a combination of

time and frequency division for spectral sharing within a cell. However, there are still 2G and 3G systems in operation that use code-division multiple access based on direct sequence spread spectrum. Channels assigned to users within a cell may be orthogonal or non-orthogonal. In the latter case the channels of two users within a cell will overlap in time, frequency, code or spatial dimensions, which is referred to as intracell interference. More details on multiple access techniques and their performance analysis will be given in Chapters 13 and 14.

Efficient cellular network designs are *interference limited* – that is, interference from within and outside a cell dominates the noise floor, since otherwise more users could be added to the system. As a result, any technique to reduce interference in cellular networks leads directly to an increase in system capacity and performance. Some methods for interference reduction in cellular networks include cell sectorization, power control, directional antennas and antenna-array signal processing, multiuser detection and interference cancellation, base station cooperation, and user scheduling. Details of these techniques will be given in Chapters 14 and 15.

1.2.3 Satellite Systems

Satellite systems are another major component of the wireless communications infrastructure [9, 10]. There is a plethora of satellite communications services available today, including mobile service to airplanes, ships, vehicles, and hand-held terminals, fixed satellite service to earth stations, as well as broadcast radio and television service. Fig. 1.6 illustrates the satellite architecture supporting all such services. One of the biggest advantages of satellites over terrestrial wireless systems is their ability to provide coverage in remote areas.

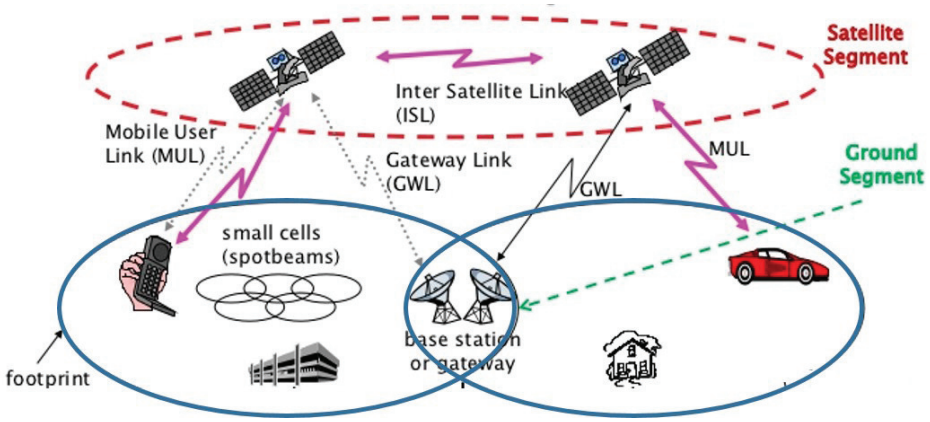


Figure 1.6: Satellite System Architecture

Satellite communication systems consist of one or more satellites communicating with stations on earth or in the air. Communication satellites typically serve as relays between stations on the ground and in some cases provide direct links to other satellites as well. Ground stations may be at fixed locations, in which case their antennas may be large to maximize received signal strength. Mobile stations and indoor fixed stations have smaller antennas, typically on the order of 1-3 meters. The coverage area or *footprint* of a satellite depends on its orbit. A GEO satellite can cover one or more continents, hence only a handful are needed to provide coverage for the entire globe. However, the transmission latency between the earth and a GEO satellite is high, on the order of 300 ms. MEO satellites have a coverage area of around ten thousand square kilometers and a latency of around 80 ms. A LEO coverage area is about a thousand square kilometers with a latency of around 10 ms.

A GEO satellite provides continuous coverage to any station within its footprint, whereas a satellite with a different orbit will have a fixed station within its footprint for only part of a 24 hour period. GEO satellites are heavy, sophisticated, and costly to build and launch. Portable stations or handsets for GEOs tend to be large and bulky due to the power required to reach the satellite. In addition, the large round-trip propagation delay to a GEO

satellite is quite noticeable in two-way voice communication. The most common service offered by GEO satellites is radio and TV broadcasting since delay is not a constraint and the satellite construction, launch, and operational costs are amortized over the many users within the GEO footprint. GEO satellites are also commonly used as a backbone link for terrestrial networks, for airplane and maritime communications, and for connectivity in remote locations.

LEO satellites are much lighter and lower in cost than GEOs to manufacture, launch, and operate. In addition, the stations and handsets in a LEO system have smaller size, transmit power, and latency than GEO systems due to the closer proximity of LEO satellites to the earth. Hence, most mobile services use LEO satellites, typically in a constellation of dozens to hundreds of satellites whose total footprint covers the locations around the globe supported by the system. Continuous coverage of a given location in a LEO system requires handoff between satellites as a fixed location on earth is within the footprint of a given LEO satellite for only 10-40 minutes. Hence, when the footprint of a LEO satellite moves away from a given station or handset, its connection is handed off to another LEO satellite so as to maintain continuous connectivity. Sophisticated antennas on LEO satellites can create very small spotbeams for mobile terminals that focus that transmission energy within a small footprint. These spotbeams allow for frequency reuse similar to that of cellular networks.

For the reasons outlined in the previous paragraph, LEOs tend to support the highest data rates, best performance, and lowest cost among satellite services. As expected, MEO satellite systems provide a compromise between the system requirements, performance, and costs compared with those of GEO and LEO systems. MEO satellite services mainly compete with GEO systems by offering better performance at a lower cost for users in relatively remote locations that are not well served by LEO or terrestrial systems.

1.2.4 Fixed Wireless Access

Fixed wireless access (FWA) systems support wireless communications between a fixed access point and multiple terminals. The FWA system architecture is shown in Fig. 1.7. The access point transmits to the receivers of multiple terminals in the downlink direction, and receives signals from the transmitters of multiple terminals in the uplink direction. Different multiple access techniques can be used to share the system spectrum between the multiple terminals.

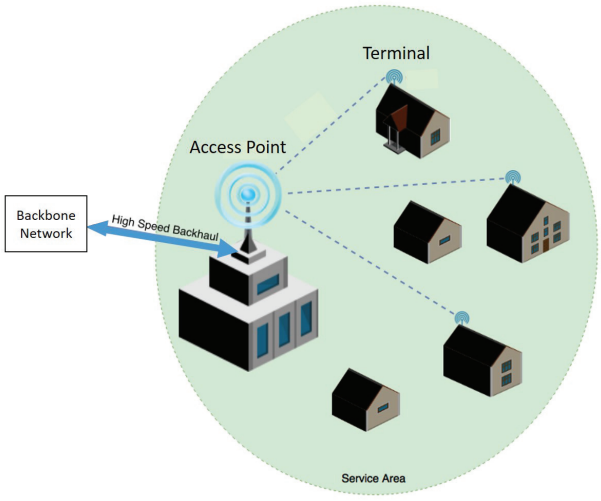


Figure 1.7: Fixed Wireless Access System

FWA systems provide an alternative to the wired broadband options of DSL, cable and fiber, whose availability

may be limited in rural areas. In the United States, two frequency bands were set aside for these systems in the late 1990s: part of the 28 GHz spectrum for local distribution systems (local multipoint distribution service, LMDS) and a band in the 2 GHz spectrum for metropolitan distribution service (multichannel multipoint distribution services, MMDS). MMDS systems were never widely deployed, and are used today in sparsely populated rural areas, where laying cables is not economically viable. The initial LMDS systems were expensive and performed poorly due to the challenges of providing service at such high frequencies. As a result these systems did not succeed commercially and were discontinued. The LMDS spectrum is just below the millimeter wave band of 30-300 GHz. Recent interest in utilizing the large amount of unregulated millimeter wave spectrum has led to advances in system, circuit, and antenna design at these high frequencies, which in turn has renewed interest in FWA systems using the LMDS and millimeter wave bands.

One of the main growth drivers of FWA was the IEEE 802.16 (WiMAX) standard, finalized in 2001. The first WiMAX standard defined operation between 2 GHz and 11 GHz for non-line-of-sight links and between 10 GHz and 66 GHz for line-of-sight links, however WiMAX systems generally operate below 6 GHz. The first commercial WiMAX systems followed the 802.16d standard with data rates of around 40 Mbps. A later version, 802.16e, was developed in the mid-2000s to support mobile users with a system design similar to that of cellular. The 802.16e systems offered 15 Mbps data rates, much higher than that of 3G cellular networks. The significantly higher data rates of 802.16e over 3G led to speculation that it would be adopted for 4G cellular. However, in the end 4G cellular adopted the LTE standard, which ultimately pushed WiMAX systems out of most mobile service markets. WiMAX is still used to support mobile services in areas where LTE has not been deployed, as well as in industrial markets such as aviation, utilities, and transportation. FWA systems based on the LTE and 5G cellular standards have been deployed to provide high-speed connectivity for homes, apartments, and office buildings.

1.2.5 Ad Hoc Wireless Networks

An ad hoc wireless network is a collection of wireless nodes that self-configure to form a network without the aid of any established infrastructure. In general, some node pairs in the network can communicate directly while others may require multihop routing through an intermediate node. Different ad hoc wireless networks may coexist in the same geographical area, as shown in Figure 1.8. Without an inherent infrastructure, the network nodes handle the necessary control and networking tasks by themselves. The nodes in an ad hoc wireless network are typically heterogeneous, and some may be gateways or routers to other networks. For example, a vehicular network may have a gateway to the backbone Internet to obtain traffic and weather information, while a mesh Wi-Fi network includes multiple access points connected together by wireless links that collectively serve all of the network's Wi-Fi devices. The Merriam-Webster dictionary lists two relevant definitions for *ad hoc*: "formed or used for specific or immediate problems", and "fashioned from whatever is immediately available". These definitions capture two of the main benefits of ad hoc wireless networks: they can be tailored to specific applications, such as the vehicle or indoor networks shown in Fig. 1.8, and they can be formed from whatever network nodes are available. Ad hoc wireless networks have other appealing features as well. They avoid the cost, installation, and maintenance of network infrastructure. They can be rapidly deployed and reconfigured. They also exhibit great robustness owing to their distributed nature, node redundancy, and the lack of single points of failure. These characteristics are especially important for military applications, and much of the groundbreaking research in ad hoc wireless networking was supported by the (Defense) Advanced Research Projects Agency (DARPA) and the U.S. Navy [11, 12]. Many of the fundamental design principles for ad hoc wireless networks were identified and investigated in that early research. However, despite many advances over the last several decades in wireless communications in general – and in ad hoc wireless networks in particular – the optimal design, performance, and fundamental capabilities of these networks remain poorly understood, at least in comparison with other wireless network paradigms. The design principles, protocols, capacity limits, and primary applications of wireless ad hoc networks are described in Chapter 16.

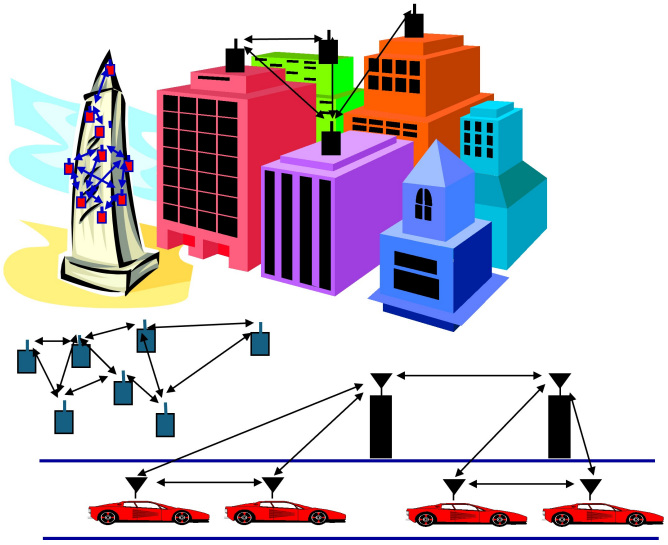


Figure 1.8: Ad Hoc Wireless Networks

As radios decrease their cost and power consumption, it becomes feasible to embed them into more types of electronic devices, which enables applications such as smart homes, sensor networks, and other interconnected systems. The most common radio standards that have emerged to support this trend are Bluetooth, ZigBee, and Z-Wave. All of these radios support a *multihop routing protocol*, hence they can form an ad hoc wireless network. In principle a multihop routing protocol can support hundreds or even thousands of nodes, but network performance generally degrades as the number of nodes increases. The Bluetooth, Zigbee, and Z-Wave short range radios have not yet demonstrated that their multihop routing protocols are feasible in practice for large numbers of nodes. More details about these standards can be found in Appendix D.3.

1.3 Wireless Spectrum

The wireless spectrum refers to the electromagnetic frequencies used for transmitting information. In principle it ranges over all frequencies, but most wireless communication systems operate in the wireless spectrum between 30 MHz and 30 GHz, with some in the millimeter wave frequency band (30 GHz-300 GHz) as well. For communication systems at frequencies below 800 MHz, antenna sizes are too large for small devices, and at frequencies above 3 GHz, signal attenuation with distance precludes long-range transmission. Given these tradeoffs, the primary frequency bands that fueled the extraordinary growth of cellular, Wi-Fi, and short-range radios like Bluetooth were in the 0.8-3 GHz range. As those frequency bands became congested, these systems moved to adjacent bands (0.6-0.8 for cellular and 3-5 GHz for cellular and Wi-Fi). Cellular and Wi-Fi systems started to expand into the millimeter wave bands as well due to the plentiful spectrum there. A similar evolution into higher frequency bands occurred in satellite systems. In particular, early satellite systems used the 4-8 GHz C band, but over time moved into the X (8-12 GHz), Ku (12-18 GHz), K (18-26 GHz) and Ka (26-40 GHz) bands.[1] The satellite frequency bands are also used for some terrestrial systems as well, either through a dedicated allocation or via protocols to share the spectrum between the satellite and ground-based system. Note that the required antenna size for good reception is inversely proportional to the signal frequency, so moving systems to a higher frequency allows for more compact antennas. However, received signal power with nondirectional antennas is proportional to the inverse of frequency

[1]The frequency ranges defining these bands differ slightly for different satellite systems

squared, so it is harder to cover large distances with high-frequency signals.

1.3.1 Licensed and Unlicensed Spectral Bands

Wireless spectrum use is regulated and either requires a license or must be used in a the set of spectral bands set aside for unlicensed use. Most countries have government agencies responsible for allocating and controlling use of the radio spectrum. In the United States, spectrum is allocated by the Federal Communications Commission (FCC) for commercial use and by the Office of Spectral Management (OSM) under the auspices of the National Telecommunications and Information Administration (NTIA) for government use. Countries throughout the world have similar regulatory agencies to regulate spectrum within their borders. Certain regions of the world have a common regulatory agency for spectrum allocation, e.g. commercial spectral allocation across Europe is governed by the European Telecommunications Standards Institute (ETSI). Satellite systems cover large areas spanning many countries and sometimes the globe. Globally, spectrum is allocated by the International Telecommunications Union Radio Communications group (ITU-R) through its World Radiocommunication Conferences (WRC). The standards arm of this body, ITU-T, adopts telecommunication standards for global systems that must interoperate across national boundaries. Regulatory agencies typically have many competing considerations in deciding how to allocate any given block of spectrum, including whether to allocate it for commercial, military, or shared use. These decisions are generally driven by a broad mandate to regulate spectrum "in the public interest," as stated in the U.S. Communications Act of 1934 establishing the FCC.

Historically government agencies allocated spectral blocks for specific uses and assigned licenses to use these blocks to specific groups or companies. For example, in the 1980s the FCC allocated spectrum around 850 MHz for analog cellular phone service, in particular 824-849 MHz for the downlink (base station to mobile) and 869-894 MHz in the uplink (mobile to base station). Spectral licenses were provided to two operators in each geographical area based on a number of criteria. While this method of licensed spectral allocation is still used in some circumstances, a fundamental shift occurred worldwide in the early 1990s for licensed spectrum to be auctioned to the highest bidder, with some restrictions in place to ensure fairness and competitive use of the spectrum. The basis for the shift was the market-based reasoning that auctions provide the fairest and most efficient way for governments to allocate the limited spectral resource and, moreover, this method provides significant revenue to the government. However, auctions are not universally supported for spectral allocation based on the contention that they can stifle innovation, limit competition, and hurt technology adoption. Specifically, the high cost of spectrum dictates that only large companies or conglomerates can purchase it. Moreover, the large investment required to obtain spectrum can delay, sometimes indefinitely, the ability to invest in infrastructure for system rollout. Finally, high spectral cost is usually passed on to the end user. The early 3G spectral auctions, with suspected collusion between bidders, low bids, and several auction winners that ultimately defaulted, provided ammunition to the opponents of spectral auctions. Lessons learned from these early auctions were adopted in the design of subsequent auctions, which generally went smoothly and raised up to tens of billions of dollars. In addition, reverse or incentive auctions were initiated starting in 2016, whereby license holders could sell back their spectrum to regulatory bodies for future auctioning. A comprehensive treatment of spectrum regulation and its allocation through auctions can be found in [13].

Spectrum is also set aside in specific frequency bands, called unlicensed bands, that are free to use without a license according to a specific set of rules. The rules may correspond to a specific access protocol to ensure fairness, restrictions on power levels, and so forth. The purpose of these *unlicensed bands* is to encourage innovation and low-cost implementation. Wi-Fi is often associated with the unlicensed frequency bands, however it is just one of the hundreds of successful unlicensed systems, which include standardized short-range radio systems as well as proprietary radios in cordless phones, wireless home security systems, baby monitors, medical equipment, inventory systems, smart meters, and keyless automobile entry systems. Indeed, it is estimated that unlicensed wireless devices contribute tens of billions of dollars annually to the US economy alone. A major difficulty of

unlicensed bands is that they can be killed by their own success. If many unlicensed devices on the same channel are used in close proximity then they interfere with each other, which can make the band unusable. Cellular networks can also operate in the unlicensed bands by following the unlicensed spectrum rules. However, since these systems can use their licensed bands for control, they have an advantage over systems such as Wi-Fi that use the unlicensed bands for both control and data transmission and can crowd out Wi-Fi users as a result [14]. Much of the spectrum above 30 GHz is unregulated or lightly regulated, and hence can be used by unlicensed users with little to no restriction.

1.3.2 Underlay Systems and Cognitive Radios

Underlay systems are another alternative for allocating spectrum. An underlay system operates as a secondary user in a licensed frequency band simultaneous with the licensed users in a manner such that the licensed users experience minimal interference from them. This is usually accomplished by spreading the signal over a very wide bandwidth, typically more than 500 MHz, and restricting its power per Hertz. The first underlay standard approved for operation was ultrawideband (UWB) communications [15]. Specifically, in 2002 the FCC approved 7500 MHz of spectrum for the use of UWB devices, with the very stringent power restriction of no more than 75 nW/MHz. This sparked regulatory activities in countries throughout Europe and Asia to also enable UWB, albeit with different restrictions than those in the US. Indeed, regulatory approval of UWB proved to be quite controversial given the complexity of characterizing how interference affects the primary band users, and the fact the UWB transmissions span many licensed users across both commercial and government domains. The regulatory challenges coupled with the severe power constraints on UWB systems ultimately proved insurmountable for most commercial systems, hence the technology failed to achieve much success [16]. The interference constraint for underlay users may alternatively be met without restricting power per Hertz by using multiple-antenna techniques to guide the underlay signals away from the spatial dimensions occupied by licensed users [17].

Following the introduction of underlay systems, regulatory bodies began exploring other innovative technologies that could make spectrum utilization in the licensed bands more flexible and efficient. This push for innovation was long overdue; other than spectral auctions and underlay systems, the basic mechanisms for licensed spectral allocation had not changed much since the inception of regulatory bodies in the early to mid-1900s. Many of the compelling ideas for exploiting technology to better utilize licensed spectrum fall under the notion of a *cognitive radio*. A cognitive radio utilizes advanced radio and signal processing technology along with novel spectrum allocation policies to support unlicensed users operating in the existing licensed spectrum, without degrading the performance of the licensed users. In particular, a cognitive radio "learns" about coexisting licensed users within its spectrum and then uses this information to utilize the spectrum without degrading the transmissions of these users [18]. Based on the nature of the coexisting user information the cognitive radio can collect, as well as a priori rules about spectrum usage, a cognitive radio seeks to *overlay* or *interweave* its signal with the transmissions of licensed nodes. Spatial underlay systems that dynamically avoid the spatial dimensions of licensed users also fall within the paradigm of cognitive radios. Cognitive radio techniques can also be applied in the unlicensed bands to reduce interference between users and thus improve spectral efficiency in these bands as well.

Interweave cognitive radios utilize unused parts of the licensed spectrum. The idea of interweave radios came about after studies conducted by the FCC and industry showed the existence of space-time-frequency voids, referred to as *spectrum holes*, in both the licensed and unlicensed bands that are not in constant use. These spectrum holes, which can change with time and geographic location, can be exploited by interweave radios to support their communication. In some cases the spectrum holes are permanent, either because a certain block of licensed spectrum is unused, or because "guard channels" between occupied channels are needed in the overall design to reduce interference between the licensed users. Such guard channels, or "white spaces," were specified in the analog television broadcast standard. As analog television bands transitioned to digital, arguments were made that such guard bands were no longer needed, and hence these white spaces could be freed up for other uses. Despite

resistance from the television broadcasters, unlicensed use of television white spaces was approved by regulatory bodies in multiple countries starting in 2010. White space devices must generally consult a database of available spectrum holes in their geographical area before using them, and such databases are relatively static. A more sophisticated interweave radio looks for dynamic spectrum holes by periodically monitoring the radio spectrum, detecting spectrum holes as they occur in time, space, and/or frequency, and then using such holes opportunistically to communicate over them. Such opportunistic use of white spaces, which was the original motivation behind the notion of cognitive radio [19], has yet to be approved for spectrum allocation.

In overlay systems the cognitive transmitter has some knowledge about the transmissions of noncognitive users in the band. This knowledge can be exploited in two ways: to mitigate the interference caused to licensed users and to improve the performance of licensed users by amplifying (relaying) their signals. In particular, an overlay user with knowledge of the licensed user's transmitted signal can use a special type of signal encoding to completely cancel the interference caused by this licensed user at its own receiver. While this coded signal will cause interference to the licensed user, by also using part of its power to amplify the licensed user's signal, the impact of this interference can be completely mitigated and, in fact, the licensed user might even experience better performance than without the existence of the overlay user. The overlay cognitive radio paradigm was originally proposed in [20] and capacity limits as well as practical implementations of overlay systems have been extensively studied since then. However, regulatory agencies have not yet considered allowing overlay systems to operate in the licensed spectrum.

Overlay, underlay, and interweave radio innovations could make spectrum utilization far more efficient than it is today, in both the licensed and unlicensed bands, thereby enabling new wireless systems, products, and applications. More details on the technologies behind these cognitive radio paradigms will be discussed in Chapter 16.7. However, even once these and other cognitive radio technologies are fully developed, changing regulatory policies to include these novel mechanisms will likely be fraught with conflicting opinions from government, industry, and the end users about what best serves the public interest in the allocation of spectrum.

1.3.3 Spectrum Allocations

As discussed in Section 1.3.1, spectrum is allocated either in licensed bands (which regulatory bodies assign to specific operators) or in unlicensed bands (which can be used by any system subject to certain operational requirements). Often different countries try to match their frequency bands for licensed and unlicensed use so that technology developed for that spectrum is compatible worldwide, however that isn't possible if a country has allocated one of these frequency bands to another use. Figure 1.9 shows the unlicensed spectrum allocations in the United States. In most cases there are similar frequency allocations in Europe and Asia. ISM Band I at 900 MHz has limited spectrum availability as well as licensed users transmitting at high power who interfere with the unlicensed users. Since performance in this band is somewhat poor, it is not heavily utilized by unlicensed devices. The U-NII bands have a total of 555 MHz of spectrum in four separate bands, with different power restrictions. In addition, some countries have imposed restrictions on parts of the U-NII bands so that unlicensed systems must avoid interference with radar systems licensed to operate in those bands.

Figure 1.10 shows the frequency bands allocated to current and legacy cellular networks in the United States along with their band name abbreviations[2]. Note from this table that 4G (LTE) systems operate in the 3.5 GHz CBRS band and the 5.2 GHz U-NII band, which are unlicensed bands, and hence must be shared with other unlicensed users. Prior to 4G, cellular networks operated only in dedicated licensed spectrum and hence interference was centrally managed. Under this new paradigm of spectrum sharing, the licensed and unlicensed users using the same band must manage shared access and interference between the different systems.

[2]Specialized Mobile Radio (SMR); Personal Communication System (PCS); Advanced Wireless Services (AWS); Digital Dividend (DD); Wireless Communications Services (WCS); Broadband Radio Service (BRS); Citizen's Broadband Radio Service (CBRS).

Band Name	Primary Use	Frequency Range (GHz)	Total Bandwidth (MHz)	Transmit Power Restriction (W)
ISM band I	Cordless phones, Z-Wave, First-generation WLANs	0.902–0.928	26	1
ISM band II	Cordless phones, Bluetooth, Zigbie, 802.11b/g	2.4–2.4835	83.5	1
ISM band III	Wi-Fi 6, 802.11a/n/ac	5.725-5.875	150	1
U-NII band I	Wi-Fi 6, 802.11a/n/ac	5.15–5.25 GHz	100	0.05
U-NII band II	Wi-Fi 6, 802.11a/n/ac	5.25–5.35 GHz	100	0.25
U-NII band III	Wi-Fi 6, 802.11a/n/ac	5.47–5.725 GHz	255	0.25
U-NII band IV	Wi-Fi 6, 802.11a/n/ac	5.725–5.825 GHz	100	1

Figure 1.9: Unlicensed Frequency Bands

	2G Systems			3G Systems			4G Systems									5G Systems			
Approximate Carrier Frequency (MHz)	800	850	1900	850	1700 2100	1900	600	700	850	1700 2100	1900	2300	2500	3500	5200	600	2500	28000	39000
Band Acronym	SMR	Cellular	PCS	Cellular	AWS	PCS	DD	SMH	Cellular	AWS	PCS	WCS	BRS	CBRS	U-NII-1	DD	BRS	K	Ka
Total Bandwidth (MHz)	14	70	200	70	90	200	70	88	70	90	200	30	70	150	100	70	140	850	3000

Figure 1.10: Cellular Frequency Band Allocations in the US and their Band Abbreviations: 2G-5G Systems.

The licensed and unlicensed frequencies allocated to different wireless systems are constantly evolving to meet their capacity demands. For example, once the unlicensed spectrum in the 900 MHz and 2.5 GHz spectrum opened up, many new products and services were launched to exploit this spectrum. This led to the opening of the 3 U-NII bands for unlicensed use. TV broadcasters that were allocated spectrum in the 700 MHz band back in the 20th century were gradually moved out of these bands to open up this spectrum for cellular networks. cellular networks are now operating in several unlicensed bands and may move into more of them. Millimeter wave spectrum above the 39 GHz Ka band is also being considered for 5G cellular networks.

1.4 Standards

Communication systems that interact with each other require standardization. Standards are typically decided on by national or international committees; in the United States this role is played by the Telecommunications Industry Association (TIA) while ETSI plays this role in Europe. The IEEE is the major player for WLAN standards development through its 802.11 working group. Other IEEE 802 working groups develop standards for wireless networks, such as short-range or fixed-wireless access networks, to complement WLANs. Cellular network standards are primarily developed by 3GPP. Standards groups typically follow a lengthy process for standards development that entails input from companies and other interested parties as well as a long and detailed review process. The standards process is a large time investment, but companies participate because incorporating their ideas into the standard gives them an advantage in developing the resulting system. In general, standards do not include all the details of the system design, rather only those needed for interoperability. This allows companies to innovate and differentiate their products from other standardized systems. The main goal of standardization is enabling systems to interoperate.

In addition to ensuring interoperability, standards also allow economies of scale and pressure prices lower. For example, WLANs typically operate in the unlicensed spectral bands, so they are not required to follow a specific standard. The first generation of WLANs were not standardized and so specialized components were needed for many systems, leading to excessively high cost that, when coupled with poor performance, led to limited adoption.

This experience resulted in a strong push to standardize the next WLAN generation, which yielded the highly successful IEEE 802.11 family of standards. Adherence of products to the 802.11 standard is certified by the Wi-Fi Alliance.

There are, of course, disadvantages to standardization. The standards process is not perfect, as company participants often have their own agendas, which do not always coincide with the best technology or the best interests of consumers. In addition, the standards process must be completed at some point, after which it becomes more difficult to add new innovations and improvements to an existing standard. Finally, the standards process can become quite politicized. This happened in Europe with 1st generation cellular networks, where each country had its own standard, and with the second generation of cellular phones in the United States which ultimately adopted three different standards. The formation of the 3GPP standards body to create a single unified cellular standard throughout much of the world was a response to these pitfalls in earlier cellular standards, and is largely responsible for the massive growth in and success of cellular technology. Hence, despite its flaws, standardization is often an essential component of wireless system design and operation in order to ensure its success.

1.5 Wireless Vision

Wireless communication is ubiquitous in the world we live in, enabling vast connectivity among people and devices as well as rapid access to information. Wireless technology impacts every aspect of modern life: culture, business, politics, economics, health, entertainment, and education. So what might the future bring?

Demand for higher data rates seems unending, hence future wireless systems could support peak speeds of hundreds or perhaps thousands of Gigabits per second. There are billions of people and locations throughout the world today without wireless (or wired) connectivity, so perhaps future systems will fill in these coverage holes so that no person or place on the planet lacks wireless connectivity. Wireless devices might shrink to such small sizes that they can be deployed within any object or living being. Some wireless devices may be powered from a very small battery or even self-powered through energy harvesting or wireless charging, eliminating the need to ever plug in. Wireless technology might evolve to support the "five-nines" standard in reliability for tradition telephone service, meaning that the service is reliable 99.999% of the time in any location, indoors and out. Finally, wireless systems must be extremely secure against natural impairments as well as eavesdroppers, attackers and spoofers.

If this vision of wireless technology comes to pass, what will it enable? In addition to providing people with voice, high-speed data, and broadcast entertainment, future wireless networks will also support machine-to-machine communications for tens of billions of devices. In the home these networks will enable a new class of intelligent electronic devices that can interact with each other and with the Internet. Such "smart" homes will drastically improve energy efficiency, security, emergency response, as well as help the elderly and disabled with assisted living. Other applications of these networks include sensing and data collection in the power grid to improve robustness and efficiency, "smart cities" that provide services such as trash collection and road maintenance when the need is detected, and in-body communications for medical devices, biosensors, and targeted drug delivery. Wireless video and virtual reality will permeate the home and any place that people congregate with entertainment, and also enable remote classrooms, remote training facilities, and remote hospitals anywhere in the world. Wireless sensor networks will improve monitoring of and response to fire hazards, toxic waste sites, stress and strain in buildings and bridges, carbon dioxide movement, and the spread of chemicals and gases at disaster sites. Finally, wireless networks with very low latency will enable distributed control systems with remote devices, sensors, and actuators linked together via wireless communication channels. Such systems will in turn enable intelligent transportation systems including self-driving vehicles, mobile robots and drones, as well as easily reconfigurable industrial automation.

The exponential growth of smartphone use and wireless Internet access has led to great optimism about wireless technology in general. Obviously not all wireless applications will flourish. While many wireless systems

and companies have enjoyed spectacular success, there have also been many failures along the way, including the first generation of wireless LANs and LEO satellite systems, as well as wide area data services, and fixed wireless access to the home. Indeed, it is impossible to predict what wireless failures and triumphs lie on the horizon. Moreover, there must be sufficient flexibility and creativity among both engineers and regulators to allow for accidental successes. It is clear, however, that the current and emerging wireless systems of today – coupled with the vision of applications that wireless can enable – ensure a bright future for wireless technology.

1.6 Technical Challenges

This section provides an overview of the many technical challenges that must be addressed to make the wireless vision a reality. These challenges extend across all aspects of the system, including hardware design, channel characterization, physical layer and multiple access techniques as well as networking protocols and architectures. Techniques to address many of these challenges are described in subsequent chapters of the book.

The design of wireless systems begins with a model for the underlying channel through which the signals will propagate. In all such channels, signal power decreases with distance due to the physics of propagation as well as attenuation from blocking objects. These signal propagation models are developed in Chapter 2. If the transmitter, receiver, or surrounding objects are moving, the channel changes randomly with time due to changing reflections and attenuation. These random channel variations, whose statistical models are developed in Chapter 3, make it difficult to design reliable systems with guaranteed performance. Channel characteristics, including signal attenuation, also depend on the frequency of operation. In particular, received power generally decreases with the carrier frequency and, in the case of free space propagation with omnidirectional antennas, it is inversely proportional to the square of this frequency. Thus, most wireless systems today operate at carrier frequencies below 5 GHz to ensure good coverage, leading to a spectrum shortage in this range of frequencies. Moving to higher carrier frequencies, such as millimeter wave (30-300 GHz) or terahertz (.3-3 THz), provides much more spectrum than what is available in the lower frequency bands. However, these higher frequencies of operation reduce range unless energy is directionally focused using multiple or directional antenna techniques. Signal propagation characteristics at these frequencies create challenges in designing reliable communication links, as described in Chapter 2.9.5. In addition, hardware components are expensive and power hungry.

The maximum data rate that can be reliably sent over a wireless (or wireline) channel is its Shannon capacity, which is derived in Chapter 4. This capacity is directly proportional to the channel bandwidth, i.e. the amount of spectrum allocated to the channel. This rate also depends on the number of antennas at the transmitter and receiver, as multiple-input multiple-output (MIMO) techniques allow for independent data streams to be transmitted along the independent spatial dimensions these multiple antennas create, as described in Chapter 10. Achievable data rates for a given wireless system also depend on signal propagation and interference characteristics. Due to these challenges, data rates for both cellular and Wi-Fi systems are several orders of magnitude lower than for a fiber optic cable, but that could change as millimeter wave systems are deployed.

In terms of hardware challenges, as the size of wireless devices shrink, breakthroughs are needed to make both the analog and digital circuitry significantly smaller, cheaper, and more energy efficient. Wireless systems operating at millimeter wave and higher frequencies require cheap and reliable RF components, which remains a significant challenge. The size, power consumption, and precision of analog-to-digital converters is also becoming a bottleneck as systems move to larger bandwidths. Finally, large antenna arrays that improve signal propagation and mitigate interference require hardware innovations, including hybrid analog and digital processing, to reduce their footprint, power consumption, and cost.

The physical layer design of wireless systems today is quite sophisticated, with dense-constellation modulation (Chapters 5-6), diversity and adaptive modulation techniques to combat random signal variations (Chapters 7 and 9), powerful error-correction coding (Chapter 8), as well as techniques to combat intersymbol interference

caused by delayed signal components that arise from channel reflections (Chapters 11-13). Systems operating in rapidly changing environments require new modulation, coding, detection, and multi-antenna techniques that are robust to such changes when adaptation to them is infeasible. Machine learning may play a role in improving physical layer techniques for channel estimation, signal detection, decoding, and equalization, particularly when channels are hard to model, hard to estimate, or rapidly varying [21, 22]. In addition, algorithms that best exploit the many spatial degrees of freedom offered by large antenna arrays are needed. For energy-constrained systems, particularly those running off non-rechargeable batteries, communication schemes must be developed that can meet performance requirements while minimizing total system power consumption (for signal transmission, reception, and processing).

Multiple access techniques, developed in Chapter 14, allow users to share the same system bandwidth. This bandwidth sharing is done either through coordinated access, as in cellular networks and current Wi-Fi standards, or through distributed techniques, as used in early generations of Wi-Fi systems. Most wireless systems use access schemes that assign orthogonal time and frequency slices of the total system bandwidth to different users. For MIMO systems, independent spatial dimensions can be used as separate channels as well. Non-orthogonal access designs improve spectral efficiency by overlapping channels in the time, frequency, code, or spatial dimensions while managing the interference that results. Channel assignment is done by the base stations in cellular networks and by access points in Wi-Fi systems. Wireless access techniques are ripe for innovation, in the centralized and distributed mechanisms used to assign channels as well as in slicing up channels in a non-orthogonal manner to increase spectral efficiency with minimal interference. Machine learning is also being applied to multiple access, resource allocation, and scheduling [21].

The network architectures and protocols for future wireless systems must support a much wider range of devices and applications than exist today. For infrastructure-based systems like cellular networks (Chapter 15), this will require a more flexible architecture, with a range of cell sizes, greater densification of cells, as well as significant data processing, dynamic optimization, and resource allocation that may be centralized, decentralized, or centralized within a subsets of nodes (coined neighborhood or "fog-based" optimization). Rethinking cellular network architectures to take into account novel forms of cooperation across base stations and users to exploit rather than mitigate interference may lead to significantly higher capacity and robustness. Reduction of latency and overhead for applications with short delay-constrained messages is also needed, as are innovations in energy-efficient architectures, e.g. for systems whose backbone is powered by batteries, solar, other forms of energy harvesting, or wireless power transfer. In ad hoc wireless networks (Chapter 16), advances in cooperative techniques for transmission, reception, and relaying have the potential to significantly improve coverage, reliability, and data rates.

Seamless handoff between different wireless networks, such as Wi-Fi and cellular, or cellular and satellite, is precluded by existing protocols and addressing mechanisms, which must evolve to allow a device to continuously connect on any and all available networks. How best to utilize a multiplicity of available networks for a given device or application is an open challenge. Many wireless systems consist of predominantly wired links, with wireless as the last hop only. Hence, challenges in wired networking, such as latency and bandwidth shortages, impact wireless networks as well. Deployment of caching and computation capabilities in base stations and Wi-Fi access points give rise to new design challenges; should applications utilize these edge capabilities or rely on more powerful capabilities situated farther from the network edge, which entail more latency to access. Security is also more difficult to implement for signals traveling over wireless channels, since the airwaves are susceptible to snooping and jamming by anyone with an RF antenna.

Chapter 1 Problems

1. As storage capability increases, we can store larger and larger amounts of data on smaller and smaller storage devices. Indeed, we can envision microscopic computer chips storing teraflops of data. Suppose this data is to be transferred over some distance. Discuss the pros and cons of putting a large number of these storage devices in a truck or drone and driving or flying them to their destination rather than sending the data electronically.

2. Describe two technical advantages and disadvantages of wireless systems that use bursty data transmission rather than continuous data transmission.

3. Fiber optic cable typically exhibits a probability of bit error of $P_b = 10^{-12}$. A form of wireless modulation, DPSK, has $P_b = 1/2\overline{\gamma}$ in some wireless channels, where $\overline{\gamma}$ is the average SNR. Find the average SNR required to achieve the same P_b in the wireless channel as in the fiber optic cable. Because of this extremely high required SNR, wireless channels typically have P_b much larger than 10^{-12}.

4. Find the round-trip delay of data sent between a satellite and the earth for LEO, MEO, and GEO satellites assuming the speed of light is $3 \cdot 10^8$ m/s. If the maximum acceptable delay for a voice system is 30 ms, which of these satellite systems would be acceptable for two-way voice communication?

5. What applications might significantly increase the demand for wireless data?

6. This problem illustrates some of the economic issues facing service providers for mixed-media systems. Suppose you are a service provider with 120 kHz of bandwidth that you must allocate between voice and data users. The voice users require 20 kHz of bandwidth and the data users require 60 kHz of bandwidth. So, for example, you could allocate all of your bandwidth to voice users, resulting in six voice channels, or you could divide the bandwidth into one data channel and three voice channels, etc. Suppose further that this is a time-division system with timeslots of duration T. All voice and data call requests come in at the beginning of a timeslot, and both types of calls last T seconds. There are six independent voice users in the system: each of these users requests a voice channel with probability .2 and pays \$.20 if his call is processed. There are two independent data users in the system: each of these users requests a data channel with probability .5 and pays \$.50 if his call is processed. How should you allocate your bandwidth to maximize your expected revenue?

7. Describe three disadvantages of using a fixed wireless access systems instead of DSL or cable. Describe three scenarios where the disadvantages override the advantages.

8. cellular networks have migrated to Hetnets consisting of a mix of macrocells and small cells in order to increase system capacity and energy efficiency. Name at least three design issues that are complicated by this trend.

9. Why does minimizing the reuse distance maximize the spectral efficiency of a cellular network?

10. This problem demonstrates the capacity increase associated with a decrease in cell size. Consider a square city of 100 square kilometers. Suppose you design a cellular network for this city with square cells, where every cell (regardless of cell size) has 100 channels and so can support 100 active users. (In practice, the number of users that can be supported per cell is mostly independent of cell size as long as the propagation model and power scale appropriately.)

    (a) What is the total number of active users that your system can support for a cell size of 1 km^2?

(b) What cell size would you use if your system had to support 250,000 active users?

Now we consider some financial implications based on the fact that users do not talk continuously. Assume that Friday from 5–6 P.M. is the busiest hour for cell-phone users. During this time, the average user places a single call, and this call lasts two minutes. Your system should be designed so that subscribers need tolerate no greater than a 2% blocking probability during this peak hour. (Blocking probability is computed using the Erlang B model: $P_b = (A^C/C!)/\left(\sum_{k=0}^{c} A^k/k!\right)$, where C is the number of channels and $A = U\mu H$ for U the number of users, μ the average number of call requests per unit time per user, and H the average duration of a call [23, Chap. 3.6].

(c) How many total subscribers can be supported in the macrocell system (1-km^2 cells) and in the microcell system (with cell size from part (b))?

(d) If a base station costs $500,000, what are the base station costs for each system?

(e) If the monthly user fee in each system is $50, what will be the monthly revenue in each case? How long will it take to recoup the infrastructure (base station) cost for each system?

Bibliography

[1] G.R.M. Garratt, *The Early History of Radio: From Faraday to Marconi*, The Institution of Engineering and Technology, 1994.

[2] S. M. Cherry, "What went wrong at Ricochet?," *IEEE Spectrum*, Vol. 39, No. 3, pp. 60-61, March 2002.

[3] W. Lemstra, V. Hayes and J. Groenewegen (eds), *The Innovation Journey of Wi-Fi: The Road to Global Success,* Cambridge University Press, 2011.

[4] D. H. Ring, "Mobile Telephony - Wide Area Coverage," *Bell System Tech. Memo*, Dec. 1947.

[5] J. Engel, R. Frenkiel, and P. Porter, "High Capacity Mobile Telephone System Feasibility Studies and System Plan", FCC Proposal, 1971.

[6] V. H. McDonald, "The cellular Concept," *Bell System Tech. J*, pp. 15-49, Jan. 1979.

[7] P. Nelson, "A new breed of broadband satellites could have you living on a desert island", *Network World*, July 2014.

[8] F. Abrishamkar and Z. Siveski, "PCS global mobile satellites," *IEEE Commun. Mag.,*, pp. 132-136, Sep. 1996.

[9] J. Pelton, S. Madry, and S. Camacho-Lara (eds), *Handbook of Satellite Applications*. Springer, Cham, 2017.

[10] S. K. Sharma, S. Chatzinotas, and P.-D. Arapoglou (eds.), *Satellite Communications in the 5G Era Year,* IET Telecommunication Series, 2018

[11] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in packet radio technology," *Proc. IEEE*, pp. 1468–96, November 1978.

[12] J. Jubin and J. D. Tornow, "The DARPA packet radio network protocols," *Proc. IEEE*, pp. 21–32, January 1987.

[13] M. Cave and W. Webb, *Spectrum Management - Using the Airwaves for Maximum Social and Economic Benefit,* Cambridge University Press, 2016.

[14] F. M. Abinader et al., "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 54-61, Nov. 2014.

[15] S. Emami, *UWB Communication Systems: Conventional and 60 GHz: Principles, Design and Standards,* Springer, 2013.

[16] L. Frenzel, "What ever happened to ultrawideband wireless," *Electronic Design,* Sep. 2014.

[17] R. Zhang and Y. C. Liang, “Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks,” *IEEE J. Sel. Topics Signal Proc.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.

[18] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio,* Cambridge University Press, 2012.

[19] J. Mitola and G. Q. Maguire, ”Cognitive radio: making software radios more personal,” *IEEE Personal Commun.*, Vol. 6, No. 4, pp. 13-18, Aug. 1999.

[20] N. Devroye, P. Mitran, and V. Tarokh, “Achievable rates in Cognitive Radio Channels,” *Proc. 39th Annual Conf. Info. Sci. and Sys.*, Mar. 2005.

[21] D. Gündüz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. Murthy, M. van der Schaar, “Machine learning in the air,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2184-2199, Oct. 2019.

[22] O. Simeone, “A very brief introduction to Machine Learning with applications to communication systems,” *IEEE Trans. on Cognitive Commun. and Networking*, vol. 4, no. 4, pp. 648-664, Dec. 2018.

[23] T. S. Rappaport, *Wireless Communications - Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

Chapter 2

Path Loss, Shadowing, and Multipath

The wireless radio channel poses a severe challenge as a medium for reliable high-speed communication. Not only is it susceptible to noise, interference, and other channel impediments, but these impediments change over time in unpredictable ways as a result of user movement and environment dynamics. In this chapter we characterize the primary phenomena that affect signal propagation: path loss and shadowing, as well as signal reflection, diffraction, and scattering. Path loss characterizes how a signal's received power decreases with transmit-receive distance. It is caused by dissipation of the power radiated by the transmitter as well as by effects of the propagation channel. Path-loss models assume that path loss is the same at a given transmit–receive distance (assuming that the path-loss model does not include shadowing effects). Shadowing is the attenuation caused by obstacles between the transmitter and receiver that absorb the transmitted signal. When the attenuation is strong, the signal is blocked. The number and type of objects that cause shadowing at any given receiver location is typically unknown. Hence attenuation due to shadowing is modeled as a random parameter. Unlike path loss, shadowing does not depend on the transmit-receive distance itself but rather on the objects between the transmitter and receiver. Reflection, diffraction, and scattering are caused by a transmitted signal interacting with objects in the environment around the transmitter or receiver. The signal components that arise due to these objects are called *multipath* components. Different multipath components arrive at the receiver with different time delays and phase shifts. When the phase shifts are aligned, the multipath components add constructively; when they are not aligned, they add destructively. This constructive and destructive addition of multipath components leads to significant variations in the received signal power.

Received power variation due to path loss occurs over long distances (100-1000 m), whereas variation due to shadowing occurs over distances that are proportional to the length of the obstructing object (10-100 m in outdoor environments and less in indoor environments). Since variations in received power due to path loss and shadowing occur over relatively large distances, these variations are sometimes referred to as *large-scale propagation effects*. The received power variations due to constructive and destructive addition of multipath components occur over very short distances, on the order of the signal wavelength, since each component's phase rotates 360 degrees over that distance. Hence, power variations due to multipath are sometimes referred to as *small-scale propagation effects*. Figure 2.1 shows an example of the received-to-transmit power ratio in decibels[1] (dB) versus log distance for the combined effects of path loss, shadowing, and multipath. As indicated in the figure, received power due to path loss alone is generally modeled as decreasing linearly with respect to the log of the transmit-receive distance, with additional slow variations due to shadowing and fast variations due to multipath.

After a brief introduction to propagation and a description of our signal model, we present the simplest model for signal propagation: free-space path loss. A signal propagating between two points with no attenuation or reflection follows the free-space propagation law. We then describe the two-ray multipath model, which augments

[1] The decibel value of x is $10 \log_{10} x$.

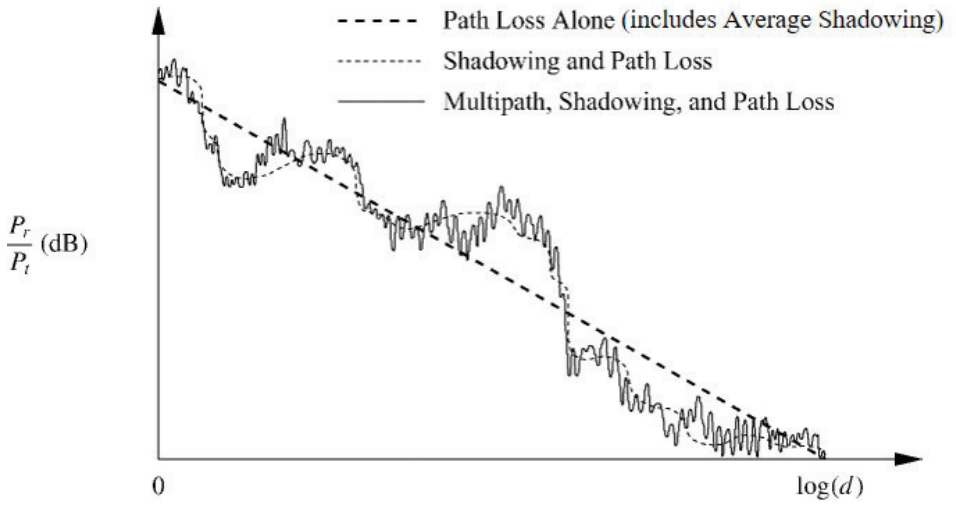


Figure 2.1: Effects of path loss, shadowing, and multipath on received power as a function of distance.

the free-space model with a single reflected ray. The two-ray model introduces the notion of a variable path-loss exponent, which gives rise to more general path loss exponent models. After introducing these models, we discuss the commonly-used log-normal model for shadowing. General ray tracing is then introduced to model the multipath components that arise due to signal reflections, diffraction, and scattering. These models approximate signal propagation according to Maxwell's equations and depend heavily on the geometry and dielectric properties of the region through which the signal propagates. If the number of multipath components is large or if the geometry and dielectric properties of the propagation environment are unknown, then statistical multipath models must be used instead of ray tracing. These statistical multipath models will be described in Chapter 3. We close the chapter by describing empirical channel models with parameters for path loss and shadowing based on measurements for both indoor and outdoor channels.

Although this chapter gives a brief overview of channel models for path loss, shadowing, and multipath, comprehensive coverage of channel and propagation models at different frequencies of interest merits a book in its own right, and in fact there are many excellent references on this topic including [1, 2, 3, 4, 5, 6]. Models specialized to multiple antenna, ultrawideband, and millimeter wave channels can be found in [7], [8], and [9], respectively.

2.1 Radio Wave Propagation

The initial understanding of radio wave propagation goes back to the pioneering work of James Clerk Maxwell, who in 1864 formulated a theory of electromagnetic propagation that predicted the existence of radio waves. In 1887, the physical existence of these waves was demonstrated by Heinrich Hertz. However, Hertz saw no practical use for radio waves, reasoning that since audio frequencies were low, where propagation was poor, radio waves could never carry voice. In 1894 Oliver Lodge used these principles to build the first wireless communication system, though its transmission distance was limited to 150 meters. By 1897 the entrepreneur Guglielmo Marconi had managed to send a radio signal from the Isle of Wight to a tugboat eighteen miles away, and in 1901 Marconi's wireless system could traverse the Atlantic ocean. These early systems used telegraph signals for communicating information. The first transmission of voice and music was made by Reginald Fessenden in 1906 using a form of amplitude modulation, which circumvented the propagation limitations at low frequencies observed by Hertz by translating signals to a higher frequency, as is done in all wireless systems today.

Electromagnetic waves propagate through environments where they are reflected, scattered, and diffracted by walls, terrain, buildings, and other objects. The ultimate details of this propagation can be obtained by solving Maxwell's equations with boundary conditions that express the physical characteristics of these obstructing objects. This often requires the calculation of the radar cross-section (RCS) of large and complex structures. Since these calculations are difficult and since the necessary parameters are often not available, approximations have been developed to characterize signal propagation without resorting to Maxwell's equations.

The most common signal propagation approximations use ray-tracing techniques based on ray-optic theory [10]. Ray-tracing approximates the propagation of electromagnetic waves by representing the wavefronts as discrete narrow beams or rays. This approximation determines the reflection and refraction effects on the wavefront but ignores the more complex scattering phenomenon predicted by Maxwell's coupled differential equations. The ray-tracing model most accurately approximates Maxwell's equations when the wavelength of the signal is much less than the size of the objects off of which it is reflected, refracted, or scattered. The simplest ray-tracing model is the two-ray model, which accurately describes signal propagation when there is one direct path between the transmitter and receiver and one reflected path. The reflected path typically bounces off the ground, and the two-ray model is a good approximation for propagation along highways or rural roads and over water. We will analyze the two-ray model in detail, as well as more complex models with additional reflected, scattered, or diffracted components. Many propagation environments are not accurately characterized by ray-tracing models. In these cases it is common to develop analytical models based on empirical measurements, and we will discuss several of the most common of these empirical models.

Often the complexity and variability of the radio channel make it difficult to obtain an accurate deterministic channel model. For these cases, statistical models are often used. The attenuation caused by signal path obstructions such as buildings or other objects is typically characterized statistically, as described in Section 2.7. Statistical models are also used to characterize the constructive and destructive interference for a large number of multipath components, as described in Chapter 3. Indoor environments tend to be less regular than outdoor environments, since the geometric and dielectric characteristics change dramatically depending on whether the indoor environment is an open factory, cubicled office, or metal machine shop. For these environments computer-aided modeling tools are available to predict signal propagation characteristics [11].

2.2 Transmit and Receive Signal Models

The transmitted and received signals in any wireless system are real-valued. The channel introduces an amplitude and phase change at each frequency of the transmitted signal so that the received signal is also real-valued. Real modulated and demodulated signals are often represented as the real part of a complex signal in order to facilitate analysis. This model gives rise to the equivalent lowpass representation of bandpass signals, which we use for our transmitted and received signals. More details on the equivalent lowpass representation of bandpass signals and systems can be found in Appendix A.

We model the transmitted signal at carrier frequency f_c as

$$\begin{aligned} s(t) &= \mathrm{Re}\{u(t)e^{j2\pi f_c t}\} \\ &= \mathrm{Re}\{u(t)\}\cos(2\pi f_c t) - \mathrm{Im}\{u(t)\}\sin(2\pi f_c t) \\ &= s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t), \end{aligned} \tag{2.1}$$

where $u(t) = s_I(t) + j s_Q(t)$ is a complex baseband signal with in-phase component $s_I(t) = \mathrm{Re}\{u(t)\}$, quadrature component $s_Q(t) = \mathrm{Im}\{u(t)\}$, bandwidth B_u, and power P_u. The signal $u(t)$ is called the *complex envelope* or *equivalent lowpass signal* of $s(t)$. We call $u(t)$ the complex envelope of $s(t)$ because the magnitude of $u(t)$ is the magnitude of $s(t)$. The phase of $u(t)$ includes any carrier phase offset. The equivalent lowpass representation of

band-pass signals with bandwidth $B \ll f_c$ allows signal manipulation via $u(t)$ irrespective of the carrier frequency. The power in the transmitted signal $s(t)$ is $P_t = P_u/2$.

For time-invariant channels, the received signal is the convolution of $s(t)$ with the channel impulse response $h(t)$ plus an additional noise component $n(t)$ introduced by the channel: $r(t) = s(t) * h(t) + n(t)$. It can be written in a similar form as the transmitted signal as:

$$r(t) = \mathrm{Re}\{v(t)e^{j2\pi f_c t}\} + n(t), \tag{2.2}$$

where $v(t)$ is the equivalent lowpass signal for the received signal without the noise. This equivalent signal depends on $h(t)$. In particular, as discussed in Appendix A, $v(t) = u(t) * c(t)$, where $c(t)$ is the equivalent lowpass channel impulse response for $h(t)$. Time-varying channels will be treated in Chapter 3.

The received signal in (2.2) consists of two terms, the first term corresponding to the transmitted signal after propagation through the channel, and the second term corresponding to the noise added by the channel. The signal-to-noise power ratio (SNR) of the received signal is defined as the power of the first term divided by the power of the second term. In this chapter (and in Chapter 3) we will neglect the random noise component $n(t)$ in our analysis, since these chapters focus on signal propagation, which is not affected by noise. However, noise will play a prominent role in the capacity and performance of wireless systems studied in later chapters.

When the transmitter or receiver is moving, the received signal will have a Doppler shift of $f_D = \frac{v}{\lambda}\cos\theta$ associated with it, where θ is the arrival angle of the received signal relative to the direction of motion, v is the receiver velocity toward the transmitter in the direction of motion, and $\lambda = c/f_c$ is the signal wavelength ($c = 3 \cdot 10^8$ m/s is the speed of light). The geometry associated with the Doppler shift is shown in Figure 2.2. The Doppler shift results from the fact that transmitter or receiver movement over a short time interval Δt causes a slight change in distance $\Delta d = v\Delta t \cos\theta$ that the transmitted signal needs to travel to the receiver. The phase change due to this path-length difference is $\Delta\phi = \frac{2\pi v}{\lambda}\Delta t \cos\theta$. The Doppler frequency is then obtained from the relationship between signal frequency and phase:

$$f_D = \frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} = \frac{v}{\lambda}\cos\theta. \tag{2.3}$$

If the receiver is moving toward the transmitter (i.e., if $-\pi/2 \leq \theta \leq \pi/2$) , then the Doppler frequency is positive; otherwise, it is negative. We will ignore the Doppler term in the path loss models of this chapter, since for typical vehicle speeds (75 km/hr) and frequencies (about 10 GHz) it is small, on the order of 1 KHz [1]. However, we will include Doppler effects in Chapter 3 on statistical fading models where it is used to characterize a random channel's rate of change.

Suppose a signal $s(t)$ of power P_t is transmitted through a given channel with corresponding received signal $r(t)$ of power P_r, where P_r is averaged over any random variations due to shadowing. We define the *linear path loss* of the channel as the ratio of transmit power to receive power:

$$P_L = \frac{P_t}{P_r}. \tag{2.4}$$

We define the *path loss* of the channel as the value of the linear path loss in decibels or, equivalently, the difference in dB between the transmitted and received signal power:

$$P_L \text{ dB} = 10\log_{10}\left(\frac{P_t}{P_r}\right) \text{ dB}. \tag{2.5}$$

The dB path loss is a nonnegative number since the channel does not contain active elements, and thus it can only attenuate the signal. The dB *path gain* is defined as the negative of the dB path loss: $P_G = -P_L = 10\ \log_{10}(P_r/P_t)$ dB. Due to the laws of physics underlying signal propagation, the dB path loss generally increases

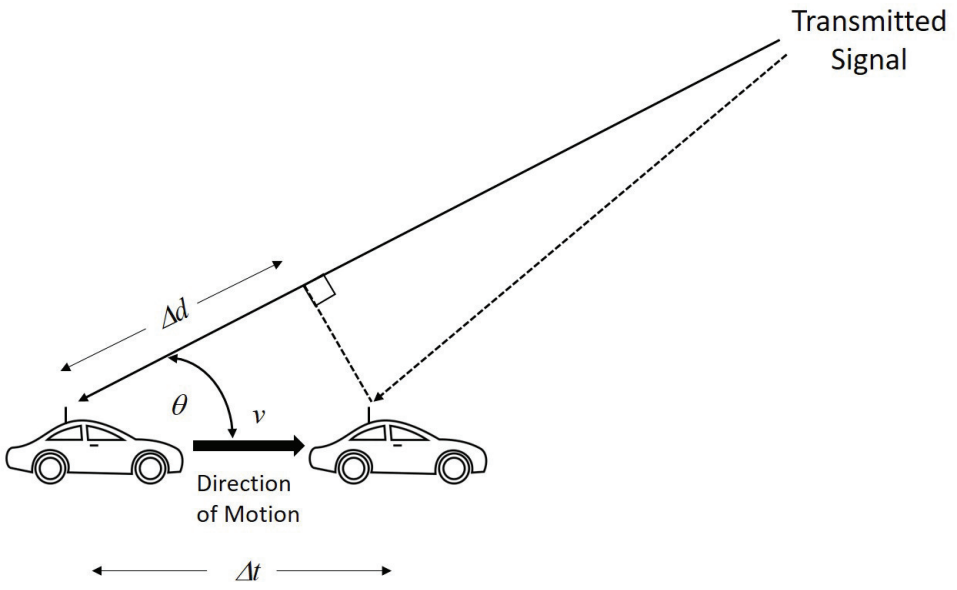


Figure 2.2: Geometry associated with Doppler shift.

with distance, as illustrated in Figure 2.1 (note that this figure shows a decreasing path gain or, equivalently, an increasing path loss). The path loss generally increases with distance, which means the received power in dB decreases with distance, as illustrated in Figure 2.1. With shadowing, also illustrated in Figure 2.1, the received power is random owing to random blockage from objects, as we discuss in Section 2.6.

2.3 Free-Space Path Loss

Consider a signal transmitted through free space to a receiver located at distance d from the transmitter. Assume there are no obstructions between the transmitter and receiver and that the signal propagates along a straight line between the two. The channel model associated with this transmission is called a line-of-sight (LOS) channel, and the corresponding received signal is called the LOS signal or ray. Under free-space path loss the received signal is given by [2]:

$$r(t) = \mathrm{Re}\left\{\left[\frac{\lambda\sqrt{G_t G_r}u(t-\tau_l)e^{-j2\pi d/\lambda}}{4\pi d}\right]e^{j2\pi f_c t}\right\}, \tag{2.6}$$

where G_t and G_r are, respectively, the transmit and receive antenna power gains in the LOS direction relative to a unity gain isotropic antenna, $\tau_l = d/c$ is the signal propagation delay of the LOS signal traveling the distance d, and the phase shift $e^{-j2\pi d/\lambda}$ is due to the distance d that the wave travels[2]. Transmit and receive directional antennas have gains G_t and G_r greater than unity in one or more angular directions relative to the idealized isotropic antenna gains $G_t = G_r = 1$. Directional antenna gains can range from 2.15 dB for a half-wavelength dipole to tens of dB in horn or dish antennas. More details on directional antenna designs and their gains can be found in [12].

[2]When the transmit and receive antennas are at the same height, the distance d equals the horizontal separation distance between the transmitter and receiver. When the transmitter and receiver are at different heights, the distance the wave travels exceeds this separation distance. These different distances will be characterized in the two-ray model (Section 2.4).

The power in the transmitted signal $s(t)$ is P_t, so from (2.6), the linear path loss for free-space propagation, a formula first introduced by Friis [13], is

$$\frac{P_r}{P_t} = G_t G_r \left[\frac{\lambda}{4\pi d}\right]^2. \tag{2.7}$$

The Friis formula (2.7) dictates that the receive power P_r equals the transmit power per unit area under free space propagation that is incident on the receive antenna, given by $P_t G_t/(4\pi d^2)$, multiplied by that antenna's *effective area* A_r, a quantity that determines how effectively the receive antenna captures this incident power. From (2.7) this effective area equals

$$A_r = \frac{\lambda^2 G_r}{4\pi}. \tag{2.8}$$

The dependence of A_r on λ^2 is due to the resonance of a signal with a given wavelength on a linear conductor half that size. Since $\lambda = c/f_c$, the Friis formulae indicates that, as the carrier frequency increases, the received power decreases as the square of the increase. Antenna arrays, discussed in Chapter 10, create a different effective area than single-element antennas, which can mitigate or remove the dependence of received power on λ [14]. The Friis formula also indicates that the received signal power falls off in inverse proportion to the square of the distance d between the transmit and receive antennas. We will see in the next section that, for other signal propagation environments, the received signal power falls off more quickly relative to this distance.

The received power can be expressed in dBm as[3]

$$P_r\ (\text{dBm}) = P_t\ (\text{dBm}) + 10\log_{10}(G_t G_r) + 20\log_{10}(\lambda) - 20\log_{10}(4\pi) - 20\log_{10}(d). \tag{2.9}$$

Equation (2.9) is an example of a *link budget equation*, which expresses the received power of a signal transmitted through a given channel (or link) as a function of the transmit power and all the losses the signal experiences on that link. *Free-space path loss* is defined as the path loss of the free-space model:

$$P_L\ (\text{dB}) = 10\log_{10}\left(\frac{P_t}{P_r}\right) = -10\log_{10}\left(G_t G_r \left[\frac{\lambda}{4\pi d}\right]^2\right). \tag{2.10}$$

The *free-space path gain* is thus

$$P_G = -P_L = 10\log_{10}\left(G_t G_r \left[\frac{\lambda}{4\pi d}\right]^2\right). \tag{2.11}$$

Example 2.1: Consider an outdoor small cell with $f_c = 2.5$ GHz, cells of radius 10 m, and isotropic antennas. Under the free-space path loss model, what transmit power is required at the base station in order for all terminals within the cell to receive a minimum power of $0.1\ \mu$W? How does this change if the system frequency is 5 GHz?

Solution: We must find a transmit power such that the terminals at the cell boundary receive the minimum required power. We obtain a formula for the required transmit power by inverting (2.7) to obtain:

$$P_t = P_r \left[\frac{4\pi d}{\sqrt{G_t G_r}\lambda}\right]^2. \tag{2.12}$$

[3]The dBm value of a power x is its dB value relative to a milliwatt: $10\log_{10}(x/.001)$. The dBW value of a power x is its dB value relative to a watt: $10\log_{10} x$, so 1W corresponds to 30 dBm.

Substituting in $G_t = G_r = 1$ (isotropic antennas), $\lambda = c/f_c = 0.12$ m, $d = 10$ m, and $P_r = 0.1$ μW yields $P_t = .1097$ W. At 5 GHz only $\lambda = .06$ changes, so $P_t = .4388$ W. We see that doubling the carrier frequency leads to a requirement for four times more transmit power, illustrating the power consumption challenges in moving wireless systems to high-frequency spectral bands.

2.4 Two-Ray Multipath Model

The two-ray model is used when a single ground (or other) reflection dominates the multipath effect, as illustrated in Figure 2.3. The received signal consists of two components: the LOS component or ray, which corresponds to the transmitted signal propagating through free space, and a reflected component or ray, which is the transmitted signal reflected off the ground. The received LOS ray is given by the free-space propagation loss formula (2.6) with the distance the signal travels set to d_0. The reflected ray shown in Figure 2.3 travels distance $d_1 = d_{11} + d_{12}$. Since the two signal components in this model travel different distances, we use d in this model and all subsequent propagation models in this chapter to denote the horizontal distance between the transmitter and receiver. The distances d_0 and d_1 then depend on d as well as the transmitter and receiver heights h_t and h_r. We see from Figure 2.3 that when $h_t = h_r$, $d = d_0$, and when d is very large relative to h_t and h_r, $d \approx d_0 \approx d_1$.

If we ignore the effect of surface wave attenuation[4] then, by superposition, the received signal for the two-ray model is

$$r_{\text{2-ray}}(t) = \text{Re}\left\{ \frac{\lambda}{4\pi} \left[\frac{\sqrt{G_0}u(t-\tau_0)e^{-j2\pi d_0/\lambda}}{d_0} + \frac{R\sqrt{G_1}u(t-\tau_1)e^{-j2\pi d_1/\lambda}}{d_1} \right] e^{j2\pi f_c t} \right\}, \tag{2.13}$$

where $\tau_0 = d_0/c$ is the time delay of the LOS ray, $\tau_1 = d_1/c$ is the time delay of the ground reflection ray, $\sqrt{G_0} = \sqrt{G_{t_0} G_{r_0}}$ is the product of the transmit and receive antenna field radiation patterns in the LOS direction, R is the ground reflection coefficient, and $\sqrt{G_1} = \sqrt{G_{t_1} G_{r_1}}$ is the product of the transmit and receive antenna field radiation patterns corresponding to the ground reflection. The *delay spread* of the two-ray model equals the difference between the delay of the LOS ray and that of the reflected ray: $(d_1 - d_0)/c$.

If the transmitted signal is narrowband relative to the delay spread such that $(\tau_1 - \tau_0 \ll B_u^{-1})$ then $u(t-\tau_1) \approx u(t-\tau_0)$. With this approximation, the received power of the two-ray model for narrowband transmission is

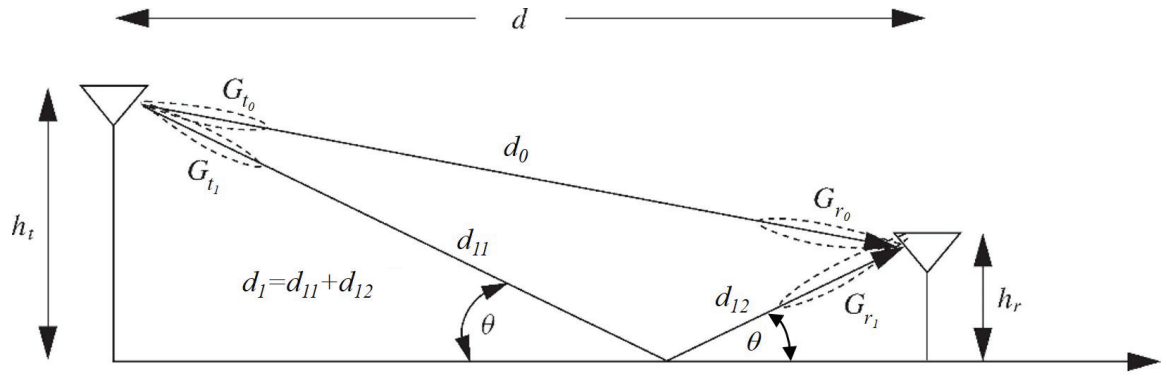


Figure 2.3: Two-ray model.

[4]This is a valid approximation for antennas located more than a few wavelengths from the ground.

$$P_r = P_t \left[\frac{\lambda}{4\pi}\right]^2 \left|\frac{\sqrt{G_0}}{d_0} + \frac{R\sqrt{G_1}e^{-j\Delta\phi}}{d_1}\right|^2, \tag{2.14}$$

where $\Delta\phi = 2\pi(d_1 - d_0)/\lambda$ is the phase difference between the two received signal components. Equation (2.14) has been shown [15] to agree closely with empirical data. From the geometry of the two-ray model, $d_1 = \sqrt{(h_t + h_r)^2 + d^2}$ and $d_0 = \sqrt{(h_t - h_r)^2 + d^2}$. Hence

$$d_1 - d_0 = \sqrt{(h_t + h_r)^2 + d^2} - \sqrt{(h_t - h_r)^2 + d^2}. \tag{2.15}$$

When d is very large compared to $h_t + h_r$, we can use a Taylor series approximation in (2.15) to get

$$\Delta\phi = \frac{2\pi}{\lambda}\left(\sqrt{(h_t + h_r)^2 + d^2} - \sqrt{(h_t - h_r)^2 + d^2}\right) \approx \frac{4\pi h_t h_r}{\lambda d}. \tag{2.16}$$

The ground reflection coefficient is given by [1, 16]

$$R = \frac{\sin\theta - Z}{\sin\theta + Z}, \tag{2.17}$$

where

$$Z = \begin{cases} \sqrt{\varepsilon_r - \cos^2\theta}/\varepsilon_r & \text{for vertical polarization,} \\ \sqrt{\varepsilon_r - \cos^2\theta} & \text{for horizontal polarization,} \end{cases} \tag{2.18}$$

and ε_r is the dielectric constant of the ground. For earth or road surfaces this dielectric constant is approximately that of a pure dielectric (for which ε_r is real with a value of about 15).

Since for asymptotically large d, $d_1 \approx d_0 \approx d$, $G_1 \approx G_0$ where we denote this approximate antenna gain as G, and $\theta \approx 0$ which by (2.17) then implies that $R \approx -1$. Substituting these approximations into (2.14) yields

$$P_r \approx P_t \left[\frac{\lambda\sqrt{G}}{4\pi d}\right]^2 |1 - e^{-j\Delta\phi}|^2. \tag{2.19}$$

Using the approximation for $\Delta\phi$ in (2.16) at large d we get the final approximation for received power at large distance d as

$$P_r \approx P_t \left[\frac{\lambda\sqrt{G}}{4\pi d}\right]^2 \left[\frac{4\pi h_t h_r}{\lambda d}\right]^2 = P_t \left[\frac{\sqrt{G}h_t h_r}{d^2}\right]^2, \tag{2.20}$$

or, in dB,

$$P_r\,\mathrm{dBm} = P_t\,\mathrm{dBm} + 10\,\log_{10}(G) + 20\,\log_{10}(h_t h_r) - 40\,\log_{10}(d). \tag{2.21}$$

Thus, in the limit of asymptotically large d, the received power falls off inversely with the fourth power of d and is independent of the wavelength λ. This is in contrast to free-space propagation where power is inversely proportion to d^2 and proportional to λ^2. The received signal power becomes independent of λ because directional antenna arrays have a received power that does not necessarily decrease with λ, and combining the direct path and reflected signal at the receiver effectively forms an antenna array.

A plot of (2.14) as a function of the transmit-receive separation distance d is shown in Figure 2.4 for $f_c = 900$ MHz, $R = -1$, $h_t = 50$ m, $h_r = 2$ m, $G_0 = G_1 = 1$, and received power normalized to 0 dB at a reference distance of 1 m. This plot reflects the case where the transmit antenna height h_t is much larger than the receiver antenna height h_r, as is typical in an urban cellular system where macrocell base stations are located on top of tall

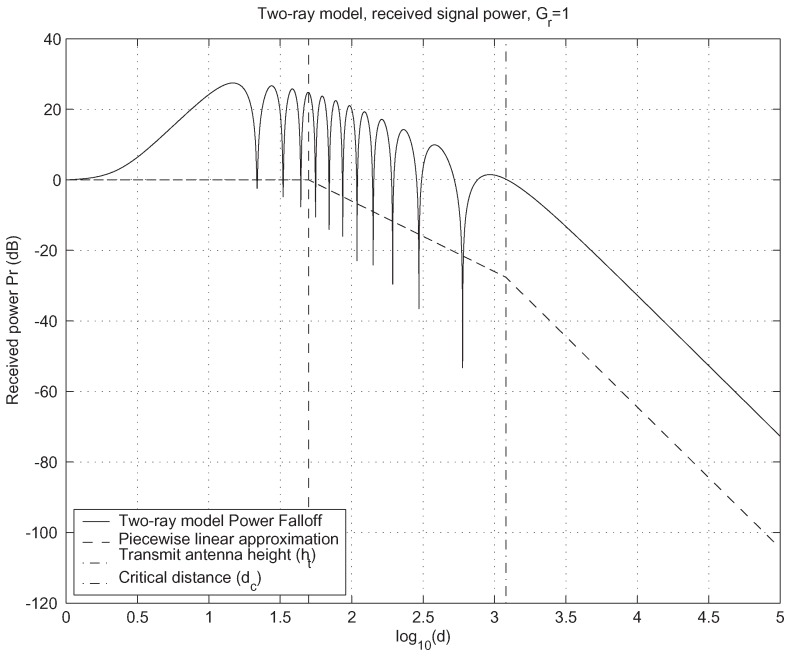


Figure 2.4: Received power versus distance for two-ray model. The received power is normalized to 0 dB at a reference distance $d_r = 1$ m (so $10\log(d/d_r) = 0$ dB for $d = 1$ m).

buildings with mobile devices near street level. The plot can be separated into three segments. For small distances $(d < h_t)$ the two rays add constructively as d increases from zero. The local maximum over this segment occurs for d such that $\Delta\phi = \pi$ in (2.16). At that distance the phases of the two multipath components are perfectly aligned ($Re^{-j\Delta\phi} = 1$). With this phase alignment the path loss over this segment is roughly proportional to $1/(d^2 + h_t^2)$ since both multipath components travel roughly the same distance $\sqrt{d^2 + h_t^2}$. That is because, for $h_r \ll h_t$,

$$d_1 = \sqrt{d^2 + (h_t + h_r)^2} \approx d_0 = \sqrt{d^2 + (h_t - h_r)^2} \approx \sqrt{d^2 + h_t^2}. \tag{2.22}$$

After this first local maximum and up to a certain critical distance d_c, the wave experiences constructive and destructive interference of the two rays as $\Delta\phi$ changes with d, resulting in a wave pattern with a sequence of maxima (when $|\Delta\phi|$ is an odd integer multiple of π) and minima (when $|\Delta\phi|$ is an even integer multiple of π). These maxima and minima are also referred to as *multipath fading*, discussed in more detail in the next chapter. At the critical distance d_c, the final maximum is reached, after which the signal power falls off proportionally with d^{-4}. This rapid falloff with distance is due to the fact that, for $d > d_c$, the signal components only combine destructively since $\Delta\phi \to 0$ as $d \to \infty$. Since the final maximum is achieved for d such that $\Delta\phi = \pi$, an approximation for d_c can be obtained by setting $\Delta\phi = \pi$ in (2.16), obtaining

$$d_c = 4h_t h_r/\lambda. \tag{2.23}$$

For the parameters used to generate Figure 2.4, $d_c = 1.2$ Km ($\log_{10} d_c = 3.08$), which is also shown in the figure.

The power falloff with distance in the two-ray model can be approximated by averaging out its local maxima and minima. This results in a piecewise linear model with three segments, which is also shown in Figure 2.4 slightly offset from the actual power falloff curve for illustration purposes. In the first segment, in order to get a constant power falloff we assume $d_0 \approx d_1 \approx h_t$, leading to a power falloff proportional to $1/h_t^2$ that is independent

of d; for distances between h_t and d_c, power falls off at -20 dB/decade; and at distances greater than d_c, power falls off at –40 dB/decade.

The critical distance d_c is proportional to the transmit antenna height h_t and to the carrier frequency $f_c = c/\lambda$. Thus, for small cells and indoor systems where h_t is ten meters or less, for f_c on the order of 1 GHz, d_c is hundreds of meters, as illustrated in Example 2.2. On the other hand, for millimeter wave systems where f_c is tens of GHz, d_c typically well exceeds a kilometer. The value of d_c can be used for system design. For example, if propagation in a cellular system obeys the two-ray model then the critical distance would be a natural size for the cell radius, since the path loss associated with interference outside the cell would be much larger than path loss for desired signals inside the cell. However, setting the cell radius to d_c could result in very large cells, as illustrated in Figure 2.4 and in the next example. Since smaller cells are more desirable – both to increase capacity and reduce transmit power – cell radii are typically much smaller than d_c. Thus, with a two-ray propagation model, power falloff within these relatively small cells goes as distance squared. Moreover, propagation in cellular systems rarely follows a two-ray model, since cancellation by reflected rays rarely occurs in all directions.

Example 2.2: Determine the critical distance for the two-ray model in an outdoor cell with $h_t = 10\ \mathrm{m},\ h_r = 3\ \mathrm{m}$ and in an indoor cell with $h_t = 3\ \mathrm{m},\ h_r = 2\ \mathrm{m}$ for $f_c = 2$ GHz.

Solution: $d_c = 4h_t h_r/\lambda = 800\ \mathrm{m}$ for the outdoor cell and $160\ \mathrm{m}$ for the indoor cell. A cell radius of $800\ \mathrm{m}$ in an outdoor cell is typical for a macrocell but a bit large for a small cell which today are on the order of 10-100 m to maintain large capacity. However, for a macrocell with an 800 m radius under these system parameters, signal power would fall off as d^2 inside the cell, while interference from neighboring cells would fall off as d^4 and thus would be greatly reduced. Similarly, 160 m is quite large for the cell radius of an indoor system, as there would typically be many walls the signal would have to penetrate for an indoor cell radius of that size. Hence an indoor system would typically have a smaller cell radius: on the order of 10–20 m, the size of one room or a few adjacent rooms in a typical building.

2.5 Path Loss Exponent Models

2.5.1 Single-Slope

The complexity of signal propagation makes it difficult to obtain a single model that characterizes path loss accurately across a range of different environments and frequencies. Accurate path loss models can be obtained from complex analytical models or empirical measurements when tight system specifications must be met or the best locations for base stations or access-point layouts must be determined. However, for general trade-off analysis of various system designs it is sometimes best to use a simple model that captures the essence of signal propagation without resorting to complicated path loss models, which are only approximations to the real channel anyway. The following single-slope model for path loss as a function of distance is a simple model that captures several important propagation characteristics,

$$P_r = P_t K \left[\frac{d_r}{d}\right]^{\gamma}, \tag{2.24}$$

Table 2.1: Typical path loss exponents

Environment	γ **range**
Urban macrocells	3.7–6.5
Urban microcells	2.7–3.5
Office building (same floor)	1.6–3.5
Office building (multiple floors)	2–6
Store	1.8–2.2
Factory	1.6–3.3
Home	3

where d_r is a reference distance for the antenna far field, γ is the path-loss exponent, and K is a unitless constant equal to the path gain P_r/P_t at distance $d = d_r$. The dB attenuation is thus

$$P_r\,(\mathrm{dBm}) = P_t\,(\mathrm{dBm}) + K\ \mathrm{dB} - 10\gamma \log_{10}(d/d_r)\,, \tag{2.25}$$

and the path loss consists of two terms, where all environment and antenna parameters are captured by the first term, and the path loss due only to distance is captured by the second term:

$$P_L\,(\mathrm{dB}) = 10\log_{10}(P_t/P_r) = -10\log_{10} K + 10\gamma \log_{10}(d/d_r). \tag{2.26}$$

The values for K, d_r, and γ can be obtained to approximate either an analytical or empirical model. In particular, the free-space path-loss model and some of the empirical models described in Section 2.7 are all of the same form as (2.25). Because of antenna near field effects at transmit distances on the order of a signal wavelength, the model (2.25) is generally valid only at transmission distances $d > d_r$, where d_r is typically assumed to be $1\ \mathrm{m}$ for systems transmitting over distances greater than this nominal value.

When the single-slope model is used to approximate empirical measurements, the value of $K < 1$ is sometimes set to the free-space path gain at distance d_r assuming isotropic antennas:

$$K\ \mathrm{dB} = 20\log_{10}\frac{\lambda}{4\pi d_r}, \tag{2.27}$$

and this assumption is supported by empirical data for free-space path loss at a transmission distance of 100 m [17]. Alternatively, K can be determined by measurement at d_r or optimized (alone or together with γ) to minimize the mean-square error (MSE) between the model and the empirical measurements [17]. The value of γ depends on the propagation environment: for propagation that approximately follows a free-space or two-ray model, γ is set to 2 or 4 respectively. The value of γ for more complex environments can be obtained via a minimum mean-square error (MMSE) fit to empirical measurements, as illustrated in Example 2.3. Alternatively, γ can be obtained from an empirically based model that takes into account frequency and antenna height [17]. Table 2.1 summarizes values of the path loss exponent γ for different environments based on data from [11, 17, 18, 19, 20, 21, 52]. Empirical measurements indicate that path loss exponents at higher frequencies tend to be higher [19, 23, 21, 22, 59, 61] whereas path loss exponents at higher antenna heights tend to be lower [17, 59, 61]. The more complex empirical models described below in Section 2.7, particularly the widely used 3GPP and WINNER II channel models [59, 61], have additional terms that explicitly capture the dependence of path loss on frequency and antenna height. Note that the wide range of empirical path-loss exponents for indoor propagation may be due to attenuation caused by floors, objects, and partitions (see Section 2.5.5).

Example 2.3: Consider the set of empirical measurements of P_r/P_t given in Table 2.2 for an indoor system at 900

MHz. Find the path loss exponent γ that minimizes the MSE between the single-slope model (2.25) and the empirical dB power measurements, assuming that $d_r = 1\,\mathrm{m}$ and K is determined from the free-space path-gain formula at this d_r. Find the received power at 150 m for the single-slope path loss model with this path loss exponent and a transmit power of 1 mW (0 dBm). Note that, since the path loss model in dB is linear, minimizing the MSE of the empirical data in dB (versus in linear units) relative to this model is a simple linear regression.

Solution: We first set up the MMSE error equation for the dB power measurements as

$$F(\gamma) = \sum_{i=1}^{5} [M_{\text{measured}}(d_i) - M_{\text{model}}(d_i)]^2, \tag{2.28}$$

where $M_{\text{measured}}(d_i)$ is the path loss measurement in Table 2.2 at distance d_i and where $M_{\text{model}}(d_i) = K - 10\gamma \log_{10}(d_i)$ is the path loss at d_i based on (2.25) since $d_r = 1$. Now using the free-space path gain formula yields $K = 20\log_{10}(.3333/(4\pi)) = -31.53\,\mathrm{dB}$. Thus

Table 2.2: path loss measurements

Distance from transmitter	$M = P_r/P_\mathrm{t}$
10 m	–70 dB
20 m	–75 dB
50 m	–90 dB
100 m	–110 dB
300 m	–125 dB

$$\begin{aligned} F(\gamma) &= (-70 + 31.53 + 10\gamma)^2 + (-75 + 31.53 + 13.01\gamma)^2 \\ &\quad + (-90 + 31.53 + 16.99\gamma)^2 + (-110 + 31.53 + 20\gamma)^2 \\ &\quad + (-125 + 31.53 + 24.77\gamma)^2 \\ &= 21682.50 - 11656.60\gamma + 1571.47\gamma^2. \end{aligned} \tag{2.29}$$

Differentiating $F(\gamma)$ relative to γ and setting it to zero yields

$$\frac{\partial F(\gamma)}{\partial \gamma} = -11656.60 + 3142.94\gamma = 0 \Rightarrow \gamma = 3.71. \tag{2.30}$$

For the received power at 150 m under the single-slope path loss model with $K = -31.53, \gamma = 3.71$, and $P_t = 0$ dBm, we have $P_r = P_t + K - 10\gamma \log_{10}(d/d_r) = 0 - 31.53 - 10 \cdot 3.71 \log_{10}(150) = -112.26$ dBm. Clearly the measurements deviate from the single-slope path loss model; this variation can be attributed to shadow fading, described in Section 2.7.

2.5.2 Multi-Slope

The multi-slope path loss model, which is piecewise linear in dB, generalizes the single-slope model by allowing for different path loss exponents at different distances. Given this greater flexibility, the multi-slope model is commonly used to analytically model empirical measurements. The multi-slope path loss model and a set of

measurement data on which it would be based is illustrated in Figure 2.5 for dB attenuation versus log distance, where the dots represent hypothetical measurements and the multi-slope model is the analytical approximation to these measurements. A multi-slope model with N segments must specify $N-1$ breakpoints $d_1, \ldots, d_{N-1}$ as well as the slopes corresponding to each segment $s_1, \ldots, s_N$. The slope of the ith segment is $s_i = -10\gamma_i$ for γ_i the path loss exponent on that segment. Different methods can be used to determine the number and location of breakpoints to be used in modeling the empirical data. Once these are fixed, the slopes corresponding to each segment can be obtained by linear regression on the data. The transition to a new slope might be instantaneous at a given breakpoint distance or entail some smoothing. Multi-slope models capture the analytical two-ray propagation model of Section 2.4. They also fit well to empirical measurements in propagation environments where the path loss exponent increases with distance [52], where the LOS path loss differs from that of non-LOS path loss [101], and in heterogeneous cellular networks where users may connect to a relatively distant macrocell base station while experiencing interference from a nearby small cell [102].

A special case of the multi-slope model is the dual-slope model. The dual-slope model is characterized by a constant path loss factor K and a path loss exponent γ_1 above some reference distance d_r and up to some breakpoint distance $d_{\rm BP}$, after which power falls off with path loss exponent γ_2. The dB path loss is thus given by:

$$P_r(d)\,(\text{dB}) = \begin{cases} P_t + K - 10\gamma_1 \log_{10}(d/d_r), & d_r \le d \le d_{\rm BP}, \\ P_t + K - 10\gamma_1 \log_{10}(d_{\rm BP}/d_r) - 10\gamma_2 \log_{10}(d/d_{\rm BP}), & d > d_{\rm BP}. \end{cases} \tag{2.31}$$

The path loss exponents, K, and $d_{\rm BP}$ are typically obtained via a regression fit to empirical data [17, 95]. The two-ray model described in Section 2.4 for $d > h_t$ can be approximated by the dual-slope model, with the breakpoint distance $d_{\rm BP} = d_c$ given by (2.23) and path loss exponents $\gamma_1 = 2$ and $\gamma_2 = 4$. Many of the measurement-based models described in Section 2.10 use the dual-slope model for path loss.

The transition between the multiple equations in the dual-slope model (2.31) can be smoothed by the following model [34, 96]:

$$P_r = \frac{P_t K}{L(d)}, \tag{2.32}$$

where

$$L(d) \triangleq \left(\frac{d}{d_r}\right)^{\gamma_1} \left(1 + \left(\frac{d}{d_{\rm BP}}\right)^{(\gamma_1-\gamma_2)q}\right)^{1/q}. \tag{2.33}$$

In this expression, q is a parameter that determines the smoothness of the path loss at the transition region close to the breakpoint distance $d_{\rm BP}$. This smoothing in this model can be extended to more than two regions [94].

2.6 Shadowing

A signal transmitted through a wireless channel will typically experience random variation due to blockage from objects in the signal path, giving rise to random variations of the received power at a given distance. Such variations are also caused by changes in reflecting surfaces and scattering objects. Thus, a model for the random attenuation due to these effects is also needed. The location, size, and dielectric properties of the blocking objects – as well as the changes in reflecting surfaces and scattering objects that cause the random attenuation – are generally unknown, so statistical models must be used to characterize this attenuation.

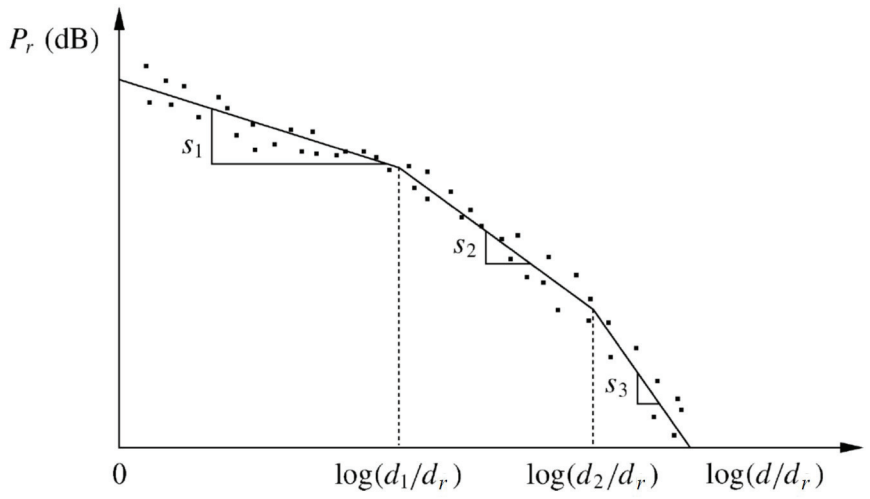


Figure 2.5: Multi-slope model for path loss.

The most common model for this additional attenuation is log-normal shadowing; under this model the dB value of the random attenuation is Gauss-distributed. The log-normal shadowing model has been empirically confirmed to model accurately the variation in received power in both outdoor and indoor radio propagation environments (see, e.g., [17, 105]). In the log-normal shadowing model, the ratio of transmit-to-receive power $\psi = P_t/P_r$ is assumed to be random with a log-normal distribution given by

$$p(\psi) = \frac{\xi}{\sqrt{2\pi}\sigma_{\psi_{\text{dB}}}\psi} \exp\left[-\frac{(10\log_{10}\psi - \mu_{\psi_{\text{dB}}})^2}{2\sigma^2_{\psi_{\text{dB}}}}\right], \ \psi > 0, \tag{2.34}$$

where $\xi = 10/\ln 10, \mu_{\psi_{\text{dB}}}$ is the mean of $\psi_{\text{dB}} = 10\log_{10}\psi$ in decibels, and $\sigma_{\psi_{\text{dB}}}$ is the standard deviation of ψ_{dB} (also in dB). The mean can be based on an analytical model or empirical measurements. For empirical measurements $\mu_{\psi_{\text{dB}}}$ equals the empirical path loss, since average attenuation from shadowing is already incorporated into the measurements. For analytical models, $\mu_{\psi_{\text{dB}}}$ must incorporate both the path loss (e.g., from a free-space or ray-tracing model) as well as average attenuation from blockage. Alternatively, path loss can be treated separately from shadowing, as described in the next section.

A random variable with a log-normal distribution is called a log-*normal random variable*. Note that if ψ is log-normal then the received power and received SNR will also be log-normal, since these are just constant multiples of ψ. For received SNR the mean and standard deviation of this log-normal random variable are also in decibels. For log-normal received power the random variable has units of power, so its mean and standard deviation will be in dBm or dBW instead of dB. The mean of ψ (the linear average path gain) can be obtained from (2.34) as

$$\mu_\psi = \mathrm{E}[\psi] = \exp\left[\frac{\mu_{\psi_{\text{dB}}}}{\xi} + \frac{\sigma^2_{\psi_{\text{dB}}}}{2\xi^2}\right]. \tag{2.35}$$

The conversion from the linear mean (in dB) to the log mean (in dB) is derived from (2.35) as

$$10\log_{10}\mu_\psi = \mu_{\psi_{\text{dB}}} + \frac{\sigma^2_{\psi_{\text{dB}}}}{2\xi}. \tag{2.36}$$

Performance in log-normal shadowing is typically parameterized by the log mean $\mu_{\psi_{\text{dB}}}$, which is referred to as the *average dB path loss* and is given in units of dB. We can compute the distribution of the dB value of ψ, $\psi_{\text{dB}} = 10\log_{10}\psi$ from the distribution of ψ given in (2.34). This computation reveals that ψ_{dB} is Gauss-distributed with mean $\mu_{\psi_{\text{dB}}}$ and standard deviation $\sigma_{\psi\text{dB}}$:

$$p(\psi_{\text{dB}}) = \frac{1}{\sqrt{2\pi}\sigma_{\psi_{\text{dB}}}} \exp\left[-\frac{(\psi_{\text{dB}} - \mu_{\psi_{\text{dB}}})^2}{2\sigma^2_{\psi_{\text{dB}}}}\right]. \tag{2.37}$$

The log-normal distribution of ψ or, equivalently, the Gaussian distribution of ψ_{dB} is defined by two parameters: $\mu_{\psi_{\text{dB}}}$ and $\sigma_{\psi_{\text{dB}}}$. Given that ψ_{dB} is based on the ratio of transmit and receive powers, it can be expressed as a difference of these powers in either dBW or dBm, as follows:

$$\psi_{\text{dB}} = 10\log_{10}(P_t/P_r) = P_t \text{ dBW} - P_r \text{ dBW} = P_t \text{ dBm} - P_r \text{ dBm}. \tag{2.38}$$

Since ψ_{dB} is Gauss-distributed, by (2.38) and properties of Gaussian random variables we have that P_r dBW is also Gauss-distributed with mean P_t dBW $- \mu_{\psi_{\text{dB}}}$ and standard deviation $\sigma_{\psi\text{dB}}$, i.e. the same standard deviation as ψ_{dB}. Similarly, P_r dBm is Gauss-distributed with mean P_t dBm $- \mu_{\psi_{\text{dB}}}$ and standard deviation $\sigma_{\psi\text{dB}}$.

The log-normal distribution is an imperfect approximation to the effect of shadowing because it has a non-zero probability that the received power exceeds the transmit power, which violates the laws of physics. In particular, since the transmit power should always exceed the received power, $\psi = P_t/P_r$ should always be greater than unity. Thus shadowing models set $\mu_{\psi_{\text{dB}}}$ to be greater than zero. Note, however, that the log-normal distribution (2.34) takes values for $0 \leq \psi \leq \infty$. Hence, for $\psi < 1$ we have $P_r > P_t$, which is physically impossible. However, this probability will be very small when $\mu_{\psi_{\text{dB}}}$ is large and positive. Thus, the log-normal model captures the underlying physical model for shadowing most accurately when $\mu_{\psi_{\text{dB}}} \gg 0$.

If the mean and standard deviation for the shadowing model are based on empirical measurements, then the question arises as to whether they should be obtained by taking averages of the linear or rather the dB values of the empirical measurements. In other words, given empirical (linear) path loss measurements $\{p_i\}_{i=1}^N$, should the mean path loss be determined as $\mu_\psi = (1/N)\sum_{i=1}^N p_i$ or as $\mu_{\psi_{\text{dB}}} = (1/N)\sum_{i=1}^N 10\log_{10} p_i$? A similar question arises for computing the empirical variance. In practice it is more common to determine mean path loss and variance based on averaging the dB values of the empirical measurements for several reasons. First, as we shall see, the mathematical justification for the log-normal model is based on dB measurements. In addition, the literature shows that obtaining empirical averages based on dB path loss measurements leads to a better approximation for the physical channel being modeled [106]. Finally, as we saw in Section 2.5.4, power falloff with distance models are often obtained by a piecewise linear approximation to empirical measurements of dB power versus the log of distance [11].

Most empirical studies for outdoor channels support a standard deviation $\sigma_{\psi_{\text{dB}}}$ ranging from 4 dB to 13 dB [1, 34, 107, 108, 109]. Larger deviations are associated with environments that contain a high density of blocking objects such as buildings and foliage outdoors, or walls and furniture indoors. Moreover, since attenuation by objects is more severe and more variable at high frequencies, and there is more variance in the number of attenuating objects at large distances, $\sigma_{\psi_{\text{dB}}}$ generally increases with both frequency and distance [75]. The mean power $\mu_{\psi_{\text{dB}}}$ depends on the path loss and building properties in the area under consideration. The mean power $\mu_{\psi_{\text{dB}}}$ varies with distance; this is due to path loss and to the fact that average attenuation from objects increases with distance owing to the potential for a larger number of attenuating objects.

The Gaussian model for the distribution of the mean received signal in dB can be justified by the following attenuation model when shadowing is dominated by the attenuation from blocking objects. The attenuation of a signal as it travels through an object of depth d is approximately equal to

$$s(d) = ce^{-\alpha d}, \tag{2.39}$$

where α is an attenuation constant that depends on the object's interior materials and associated dielectric properties while c is a fixed attenuation constant that doesn't depend on the object's depth (e.g. the penetration loss the signal experiences in entering and exiting the object). If we assume that α is approximately equal for all blocking objects and that the ith blocking object has a random depth d_i, then the attenuation of a signal as it propagates through this region is

$$s(d_t) = ce^{-\alpha \Sigma_i d_i} = ce^{-\alpha d_t}, \tag{2.40}$$

where $d_t = \sum_i d_i$ is the sum of the random object depths through which the signal travels. If there are many objects between the transmitter and receiver, then by the central limit theorem we can approximate d_t by a Gaussian random variable. Thus, $\ln s(d_t) = -\alpha d_t$ will have a Gaussian distribution with mean μ and standard deviation σ. The value of σ will depend on the environment.

Example 2.4: In Example 2.3 we found that the exponent for the single-slope path loss model that best fits the measurements in Table 2.2 was $\gamma = 3.71$. Assuming the single-slope path loss model with this exponent, $d_r = 1$m, and the same $K = -31.53$ dB, find $\sigma^2_{\psi_{\text{dB}}}$, the variance of log-normal shadowing about the mean path loss based on these empirical measurements.

Solution:

The sample variance relative to the simplified path loss model with $\gamma = 3.71$ is

$$\sigma^2_{\psi_{\text{dB}}} = \frac{1}{5}\sum_{i=1}^{5}[M_{\text{measured}}(d_i) - M_{\text{model}}(d_i)]^2, \tag{2.41}$$

where $M_{\text{measured}}(d_i)$ is the path loss measurement in Table 2.2 at distance d_i and $M_{\text{model}}(d_i) = K - 37.1\log_{10}(d_i)$ since $d_r = 1$m. This yields

$$\begin{aligned}\sigma^2_{\psi_{\text{dB}}} &= \frac{1}{5}[(-70 + 31.53 + 37.1)^2 + (-75 + 31.53 + 48.27)^2 \\ &\quad + (-90 + 31.53 + 63.03)^2 + (-110 + 31.53 + 74.2)^2 \\ &\quad + (-125 + 31.53 + 91.90)^2] \\ &= 13.28.\end{aligned} \tag{2.42}$$

Thus, the standard deviation of shadow fading on this path is $\sigma_{\psi_{\text{dB}}} = 3.64$ dB. Note that the bracketed term in the displayed expression equals the MMSE formula (2.29) from Example 2.3 with $\gamma = 3.71$.

Extensive measurements have been taken to characterize the empirical autocorrelation function of the shadow fading process over distance for different environments at different frequencies (see e.g. [108, 110, 111, 112, 113]). The most common analytical model for this function, first proposed by Gudmundson [108] and based on empirical measurements, assumes that the shadowing $\psi(d)$ is a first-order autoregressive process where the covariance between shadow fading at two points separated by distance δ is characterized by

$$A(\delta) = \mathbf{E}[(\psi_{\text{dB}}(d) - \mu_{\psi_{\text{dB}}})(\psi_{\text{dB}}(d+\delta) - \mu_{\psi_{\text{dB}}})] = \sigma^2_{\psi_{\text{dB}}}\rho_D^{\delta/D}, \tag{2.43}$$

where ρ_D is the normalized covariance between two points separated by a fixed distance D. This covariance must be obtained empirically, and it varies with the propagation environment and carrier frequency. Measurements indicate that for suburban macrocells with $f_c = 900$ MHz, $\rho_D = 0.82$ for $D = 100$ m and for urban microcells

with $f_c = 2$ GHz, $\rho_D = 0.3$ for $D = 10$ m [108, 112]. This model can be simplified and its empirical dependence removed by setting $\rho_D = 1/e$ for distance $D = X_c$, which yields

$$A(\delta) = \sigma^2_{\psi_{\text{dB}}} e^{-\delta/X_c}. \tag{2.44}$$

The *decorrelation distance* X_c in this model is the distance at which the signal autocovariance equals $1/e$ of its maximum value and is on the order of the size of the blocking objects or clusters of these objects. For outdoor systems, X_c typically ranges from 50 m to 100 m [111, 112]. For users moving at velocity v, the shadowing decorrelation in time τ is obtained by substituting $v\tau = \delta$ in (2.43) or (2.44). For a linear array of antennas, discussed in Chapter 10, shadowing can vary with the angle between the array and the signal's LOS path. Autocorrelation relative to this angle has been investigated in [110, 112].

The first-order autoregressive correlation model (2.43) and its simplified form (2.44) are easy to analyze and to simulate. Specifically, one can simulate ψ_{dB} by first generating a white Gaussian noise process with power $\sigma^2_{\psi_{\text{dB}}}$ and then passing it through a first-order filter with response $\rho_D^{\delta/D}$ for a covariance characterized by (2.43) or response $e^{-\delta/X_c}$ for a covariance characterized by (2.44). The filter output will produce a shadowing random process with the desired correlation properties [108, 109].

2.7 Combined Path Loss and Shadowing

Models for path loss and shadowing can be combined to capture power falloff versus distance along with the random attenuation about this path loss from shadowing. In this combined model, average dB path loss ($\mu_{\psi_{\text{dB}}}$) is characterized by the path loss model while shadow fading, with a mean of 0 dB, creates variations about this path loss, as illustrated by the path loss and shadowing curve in Figure 2.1. Specifically, this curve plots the combination of the single-slope path loss model (2.24) and the log-normal shadowing random process defined by (2.37) and (2.44).

2.7.1 Single-Slope Path Loss with Log-Normal Shadowing

For the combined model of single-slope path loss (2.25) and log-normal shadowing (2.36), the ratio of received to transmitted power in dB is given by

$$\frac{P_r}{P_t}\,\text{dB} = 10\log_{10} K - 10\gamma \log_{10}(d/d_r) - \psi_{\text{dB}}, \tag{2.45}$$

where ψ_{dB} is a Gauss-distributed random variable with variance $\sigma^2_{\psi_{\text{dB}}}$. The mean of ψ_{dB} is assumed to be zero when the term $10\log_{10} K$ captures average shadowing. When this is not the case, for example when K is calculated from the free-space path gain formula at $d = d_r$, then the mean of ψ_{dB} is positive and equal to the average shadowing loss over all distances. In (2.45), and as shown in Figure 2.1, the path loss increases linearly relative to $\log_{10} d/d_r$ with a slope of -10γ dB/decade, where γ is the path loss exponent. The variations due to shadowing change more rapidly, on the order of the decorrelation distance X_c.

Examples 2.3 and 2.4 illustrated the combined model for single-slope path loss and log-normal shadowing based on the measurements in Table 2.2, where path loss obeys the single-slope path loss model with $K = -31.53\,\text{dB}$ and path loss exponent $\gamma = 3.71$ and where shadowing obeys the log-normal model with mean given by the path loss model and standard deviation $\sigma_{\psi_{\text{dB}}} = 3.65\,\text{dB}$.

2.7.2 Outage Probability

The combined effects of path loss and shadowing have important implications for wireless system design. In wireless systems there is typically a target minimum received power level P_{min} or, equivalently, a minimum SNR $\text{SNR}_{\text{min}} = P_{\text{min}}/(N_0 B)$ below which performance becomes unacceptable (e.g., the caller's voice on a cell phone is unintelligible). However, with shadowing the received power at any given distance from the transmitter is log-normally distributed with some probability of falling below P_{min}. We define *outage probability* $P_{\text{out}}(P_{\text{min}}, d)$ to be the probability that the received power at a given distance d, $P_r(d)$, falls below P_{min}: $P_{\text{out}}(P_{\text{min}}, d) = p(P_r(d) < P_{\text{min}})$. For the combined path loss and shadowing model of Section 2.8 this becomes

$$p(P_r(d) \leq P_{\text{min}}) = 1 - Q\left(\frac{P_{\text{min}} - (P_t + 10\log_{10} K - 10\gamma \log_{10}(d/d_r))}{\sigma_{\psi_{\text{dB}}}}\right), \tag{2.46}$$

where the Q-function is defined as the probability that a Gaussian random variable X with mean 0 and variance 1 is greater than z:

$$Q(z) \triangleq p(X > z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \tag{2.47}$$

The conversion between the Q-function and complementary error function is

$$Q(z) = \frac{1}{2}\,\text{erfc}\left(\frac{z}{\sqrt{2}}\right). \tag{2.48}$$

We will omit the arguments P_{min} and d of P_{out} when the context is clear or in generic references to outage probability.

Example 2.5: Find the outage probability at 150 m for a channel based on the single-slope path loss and shadowing models of Examples 2.3 and 2.4, assuming a transmit power of $P_t = 10$ mW and minimum power requirement of $P_{\text{min}} = -110.5$ dBm.
Solution: We have $P_t = 10\ \text{mW} = 10\ \text{dBm}$. Hence,

$$\begin{aligned} &P_{\text{out}}(-110.5\ \text{dBm},\ 150\,\text{m}) \\ &\quad = p(P_r(150\,\text{m}) < -110.5\,\text{dBm}) \\ &\quad = 1 - Q\left(\frac{P_{\text{min}} - (P_t + 10\log_{10} K - 10\gamma \log_{10}(d/d_r))}{\sigma_{\psi_{\text{dB}}}}\right) \\ &\quad = 1 - Q\left(\frac{-110.5 - (10 - 31.53 - 37.1\log_{10}(150))}{3.65}\right) \\ &\quad = 0.0120. \end{aligned} \tag{2.49}$$

An outage probability of l% is a typical target in wireless system designs.

2.7.3 Cell Coverage Area and Percentage

The *coverage area* of a given cell in a cellular system is defined as the area of locations within the cell where the received power is above a given minimum. Consider a base station inside a circular cell of a given radius R. All

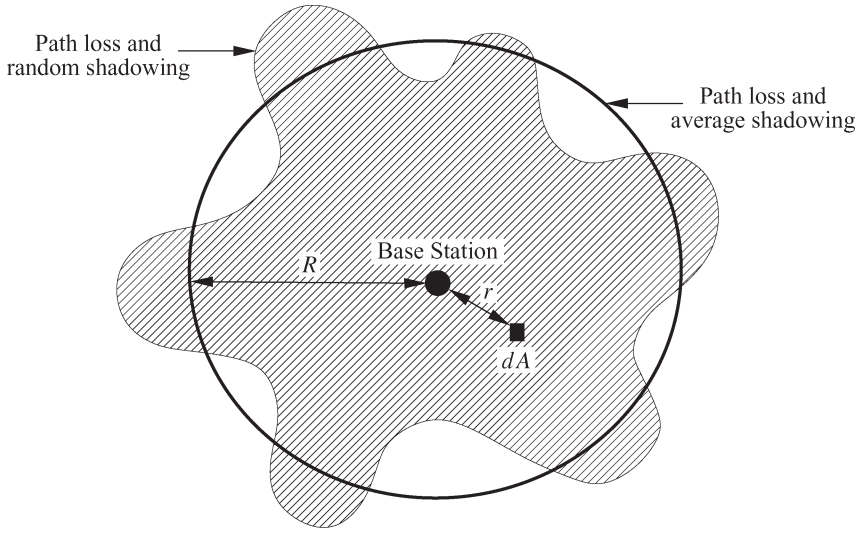


Figure 2.6: Contours of constant received power.

mobiles within the cell require some minimum received SNR for acceptable performance. Assuming a given model for noise, the SNR requirement translates to a minimum received power $P_{\min}$ throughout the cell. The transmit power at the base station is designed for an *average* received power at the cell boundary of $\overline{P}_R$, averaged over the shadowing variations. In the absence of shadowing, the coverage area of this system is πR^2 since all locations have received power above the required minimum. However, shadowing will cause some locations within the cell to have received power below $\overline{P}_R$, and others will have received power exceeding $\overline{P}_R$. This is illustrated in Figure 2.6, where we show contours of constant received power based on a fixed transmit power at the base station for path loss and average shadowing and for path loss and random shadowing. For path loss and average shadowing, constant power contours form a circle around the base station because combined path loss and average shadowing is the same at a uniform distance from the base station. For path loss and random shadowing, the contours form an amoeba-like shape due to the random shadowing variations about the average.

The constant power contours for combined path loss and random shadowing indicate the challenge that shadowing poses in cellular system design. Specifically, it is not possible for all users at the cell boundary to receive the same power level. Thus, either the base station must transmit extra power to ensure users affected by shadowing receive their minimum required power $P_{\min}$, which causes excessive interference to neighboring cells, or some users within the cell will find their minimum received power requirement unmet. In fact, since the Gaussian distribution has infinite tails, under this model *any* mobile in the cell has a nonzero probability of experiencing received power below its required minimum, even if the mobile is close to the base station. The model matches propagation scenarios in practice since a mobile may be in a tunnel or blocked by a large building, regardless of its proximity to the base station.

The *cell coverage percentage* is defined as the expected percentage of locations within a cell where received power exceeds $P_{\min}$. The cell coverage percentage under path loss and shadowing, also referred to as the fraction of useful service area, was derived by Reudink in [1, Chapter 2.5.3] as follows. The percentage of area within a cell where the received power exceeds the minimum required power $P_{\min}$ is obtained by taking an incremental area dA at radius r from the base station in the cell, as shown in Figure 2.6. Let $P_r(r)$ be the received power in dA from combined path loss and shadowing. Then the total area within the cell where the minimum power requirement is exceeded is obtained by integrating over all incremental areas where this minimum is exceeded:

$$C = \mathbf{E}\left[\frac{1}{\pi R^2}\int_{\text{cell area}} 1[P_r(r) > P_{\min} \text{ in } dA]\, dA\right]$$
$$= \frac{1}{\pi R^2}\int_{\text{cell area}} \mathbf{E}[1[P_r(r) > P_{\min} \text{ in } dA]]\, dA, \tag{2.50}$$

where $1[\cdot]$ denotes the indicator function. Define $P_A(r) = p(P_r(r) > P_{\min})$ in dA. Then $P_A(r) = \mathbf{E}[1[P_r(r) > P_{\min} \text{ in } dA]]$. Making this substitution in (2.50) and using polar coordinates for the integration yields

$$C = \frac{1}{\pi R^2}\int_{\text{cell area}} P_A(r)\, dA = \frac{1}{\pi R^2}\int_0^{2\pi}\int_0^R P_A(r) r\, dr\, d\theta. \tag{2.51}$$

The *outage probability within the cell* is defined as the percentage of area within the cell that does not meet its minimum power requirement $P_{\min}$; that is, $P_{\text{out}}^{\text{cell}} = 1 - C$. Given the log-normal distribution for the shadowing, we have

$$P_A = p(P_r(r) \geq P_{\min}) = Q\left(\frac{P_{\min} - (P_t + 10\log_{10} K - 10\gamma \log_{10}(r/d_0))}{\sigma_{\psi_{\text{dB}}}}\right)$$
$$= 1 - P_{\text{out}}(P_{\min}, r), \tag{2.52}$$

where P_{out} is the outage probability defined in (2.46) with $d = r$. Locations within the cell with received power below $P_{\min}$ are said to be *outage locations*.

Combining (2.51) and (2.52) yields[5]

$$C = \frac{2}{R^2}\int_0^R rQ\left(a + b\ln\frac{r}{R}\right) dr, \tag{2.53}$$

where

$$a = \frac{P_{\min} - \overline{P}_r(R)}{\sigma_{\psi_{\text{dB}}}}, \qquad b = \frac{10\gamma\log_{10}(e)}{\sigma_{\psi_{\text{dB}}}}, \tag{2.54}$$

and $\overline{P}_r(R) = P_t + 10\log_{10} K - 10\gamma\log_{10}(R/d_0)$ is the received power at the cell boundary (distance R from the base station) due to path loss alone. Applying integration by parts to solve this integral yields a closed-form solution for C in terms of a and b:

$$C = Q(a) + \exp\left[\frac{2-2ab}{b^2}\right] Q\left(\frac{2-ab}{b}\right). \tag{2.55}$$

If the target minimum received power equals the average power at the cell boundary, $P_{\min} = \overline{P}_r(R)$, then $a = 0$ and the cell coverage percentage simplifies to

$$C = \frac{1}{2} + \exp\left[\frac{2}{b^2}\right] Q\left(\frac{2}{b}\right). \tag{2.56}$$

Note that with this simplification C depends only on the ratio $\gamma/\sigma_{\psi_{\text{dB}}}$. Moreover, owing to the symmetry of the Gaussian distribution, under this assumption the outage probability at the cell boundary $P_{\text{out}}(\overline{P}_r(R), R) = 0.5$.

[5]Recall that (2.52) is generally valid only for $r \geq d_0$, yet to simplify the analysis we have applied the model for all r. This approximation will have little impact on cell coverage percentage, since d_0 is typically very small compared to R and the outage probability for $r < d_0$ is negligible.

Example 2.6: Find the cell coverage percentage for a cell with the combined path loss and shadowing models of Examples 2.3 and 2.4, a cell radius of 600 m, a base station transmit power of $P_t = 100\,\text{mW} = 20\,\text{dBm}$, and a minimum received power requirement of $P_{\text{min}} = -110\,\text{dBm}$ and also one of $P_{\text{min}} = -120\,\text{dBm}$.

Solution: We first consider $P_{\text{min}} = -110$ and check if $a = 0$ to see whether we should use the full formula (2.55) or the simplified formula (2.56). We have $\overline{P}_r(R) = P_t + K - 10\gamma \log_{10}(600) = 20 - 31.53 - 37.1 \log_{10}(600) = -114.6\,\text{dBm} \neq -110\,\text{dBm}$, so we use (2.55). Evaluating a and b from (2.54) yields $a = (-110 + 114.6)/3.65 = 1.26$ and $b = (37.1 \cdot 0.434)/3.65 = 4.41$. Substituting these into (2.55) yields

$$C = Q(1.26) + \exp\left[\frac{2 - 2(1.26 \cdot 4.41)}{4.41^2}\right] Q\left(\frac{2 - (1.26)(4.41)}{4.41}\right) = .59, \tag{2.57}$$

which would be a very low cell coverage percentage value for an operational cellular system (lots of unhappy customers). Now considering the less stringent received power requirement $P_{\text{min}} = -120\,\text{dBm}$ yields $a = (-120 + 114.6)/3.65 = -1.479$ and the same $b = 4.41$. Substituting these values into (2.55) yields $C = .988$, a much more acceptable value for cell coverage percentage.

Example 2.7: Consider a cellular system designed so that $P_{\text{min}} = \overline{P}_r(R)$. That is, the received power due to path loss and average shadowing at the cell boundary equals the minimum received power required for acceptable performance. Find the cell coverage percentage for path loss values $\gamma = 2, 4, 6$ and $\sigma_{\psi_{\text{dB}}} = 4, 8, 12$, and explain how cell coverage percentage changes as γ and $\sigma_{\psi_{\text{dB}}}$ increase.

Solution: For $P_{\text{min}} = \overline{P}_r(R)$ we have $a = 0$, so cell coverage percentage is given by the formula (2.56). The cell coverage percentage thus depends only on the value for $b = 10\gamma \log_{10}(e)/\sigma_{\psi_{\text{dB}}}$, which in turn depends only on the ratio $\gamma/\sigma_{\psi_{\text{dB}}}$. Table 2.3 contains cell coverage percentage evaluated from (2.56) for the different γ and $\sigma_{\psi_{\text{dB}}}$ values.

Table 2.3: cell coverage percentage for different γ and $\sigma_{\psi_{\text{dB}}}$

γ	$\sigma_{\psi_{\text{dB}}}$=4 dB	$\sigma_{\psi_{\text{dB}}}$=8 dB	$\sigma_{\psi_{\text{dB}}}$=12 dB
2	.77	.67	.63
4	.85	.77	.71
6	.90	.83	.77

Not surprisingly, for fixed γ the cell coverage percentage increases as $\sigma_{\psi_{\text{dB}}}$ decreases; this is because a smaller $\sigma_{\psi_{\text{dB}}}$ means less variation about the mean path loss. Without shadowing we have 100% cell coverage percentage (since $P_{\text{min}} = \overline{P}_r(R)$) and so we expect that, as $\sigma_{\psi_{\text{dB}}}$ decreases to zero, cell coverage percentage increases to 100%. It is a bit more puzzling that for a fixed $\sigma_{\psi_{\text{dB}}}$ the cell coverage percentage increases as γ increases, since a larger γ implies that received signal power falls off more quickly. But recall that we have set $P_{\text{min}} = \overline{P}_r(R)$, so the faster power falloff is already taken into account (i.e., we need to transmit at much higher power with $\gamma = 6$ than with $\gamma = 2$ for this equality to hold). The reason cell coverage percentage increases with path loss exponent under this assumption is that, as γ increases, the transmit power must increase to satisfy $P_{\text{min}} = \overline{P}_r(R)$. This results in higher average power throughout the cell, yielding a higher percentage of locations in the cell that have the desired minimum power.

2.8 General Ray Tracing

In a typical urban or indoor environment, a radio signal transmitted from a fixed source will encounter multiple objects in the environment, as shown in Figure 2.7. These objects produce reflected, diffracted, or scattered copies of the transmitted signal, which are not captured by the free-space and two-ray models discussed earlier. These additional copies of the transmitted signal, known as multipath components, can be attenuated in power, delayed in time, and shifted in phase and/or frequency with respect to the LOS signal path at the receiver. The transmitted signal and its multipath components are summed together at the receiver, which can produce rapid fluctuations in received signal power due to constructive and destructive combining of the components. This multipath fading was also exhibited in the two-ray model. If the arriving signal components have delay differences that exceed a symbol time, this leads to distortion in the received signal relative to the transmitted signal.

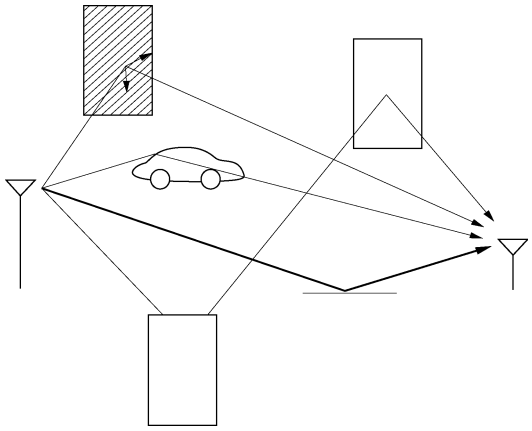


Figure 2.7: Reflected, diffracted, and scattered wave components.

In ray tracing we assume a finite number of reflectors with known location and dielectric properties. The details of the multipath propagation can then be solved using Maxwell's equations with appropriate boundary conditions. However, the computational complexity of this solution makes it impractical as a general modeling tool. Ray-tracing techniques approximate the propagation of electromagnetic waves by representing the wavefronts as simple particles. Thus, the effects of reflection, diffraction, and scattering on the wavefront are approximated using simple geometric equations instead of Maxwell's more complex wave equations. The error of the ray-tracing approximation is smallest when the receiver is many wavelengths from the nearest scatterer and when all the scatterers are large relative to a wavelength and fairly smooth. Comparison of the ray-tracing method with certain empirical data sets shows that it can accurately model received signal power in rural areas [24], along city streets when both the transmitter and receiver are close to the ground [24, 26, 27], and in indoor environments with appropriately adjusted diffraction coefficients [28].

If the transmitter, receiver, and reflectors are all immobile, then the characteristics of the multiple received signal paths are fixed. However, if the transmitter or receiver are moving then the characteristics of the multiple paths vary with time. These time variations are deterministic when the number, location, and characteristics of the reflectors are known over time. Otherwise, statistical models must be used. Similarly, if the number of reflectors is large, or if the reflector surfaces are not smooth such that each reflector generates many signal paths, then we must use statistical approximations to characterize the received signal. We will discuss statistical fading models for propagation effects in Chapter 3. Hybrid models, which combine ray tracing and statistical fading, can also be found in the literature [29, 30]; however, we will not describe them here.

The most general ray tracing model includes all attenuated, diffracted, and scattered multipath components.

This model uses all of the geometrical and dielectric properties of the objects surrounding the transmitter and receiver [37, 38, 39, 31]. Since this information is site-specific, general ray tracing methods are not used to obtain general theories about system performance and layout but rather for modeling propagation for a given transmit-receive configuration in a given environment. General ray tracing uses geometrical optics to trace the propagation of the LOS and reflected signal components as well as signal components from object diffraction and diffuse scattering. There is no limit to the number of multipath components at a given receiver location: the strength of each component is derived explicitly based on the locations and dielectric properties of the objects that generate them. In general, the LOS and reflected paths provide the dominant components of the received signal, since diffraction and scattering losses are high. However, in regions close to scattering or diffracting surfaces these other multipath components may dominate, especially if the LOS and reflected rays are blocked. Open source and commercial software for ray tracing has been developed for propagation modeling in both indoor and outdoor environments. In some of these programs, computer graphics are combined with aerial photographs (outdoor channels) or architectural drawings (indoor channels) to obtain a three-dimensional geometric picture of the environment.

The following section describes a ray tracing model capturing reflected rays only, including the commonly-used 6-ray and 10-ray models for a signal propagating along a straight street or hallway. We then present ray tracing models for signals that are reflected as well as diffracted and scattered. We also define the important parameter of local mean received power associated with these ray tracing models.

2.8.1 Multi-Ray Reflections

Most wireless channels have more than just the single reflected ray captured in the two-ray model described in Section 2.4. The multi-ray reflection model, developed in [25, 26], is a generalization of the two-ray model based on a propagation environment called a dielectric "canyon." This model captures all rays that experience one or more reflections on the path between the transmitter and receiver up to a certain maximum number of reflections.

In an outdoor setting, a dielectric canyon approximates propagation in a city with rectilinear streets[6] with buildings along both sides of the street as well as transmitter and receiver antenna heights that are close to street level. Theoretically, an infinite number of rays can be reflected off the building fronts to arrive at the receiver; in addition, rays may also be back-reflected from buildings behind the transmitter or receiver. However, since some of the signal energy is dissipated with each reflection, signal paths corresponding to more than two or three reflections can generally be ignored. When the street layout is relatively straight, back reflections are usually negligible also. Experimental data show that the dielectric canyon model of ten rays closely approximates signal propagation through cities with a rectilinear street layout [26]. The ten-ray model incorporates the LoS path as well as all nine paths with one, two, or three reflections: specifically, there is the ground-reflected, single-wall reflected, double-wall reflected, triple-wall reflected, wall-ground reflected, and ground-wall reflected paths. There are two of each type of wall-reflected path, one for each side of the street. An overhead view of the ten-ray model is shown in Figure 2.8. A simpler model, the six-ray model, includes a subset of these ten rays: the LOS, ground reflection, two single-wall reflections, and two double-wall reflections.

For the ten-ray model, the received signal is given by

$$r_{\text{10-ray}}(t) = \text{Re}\left\{\frac{\lambda}{4\pi}\left[\sum_{i=0}^{9}\frac{R_i\sqrt{G_i}u(t-\tau_i)e^{-j2\pi d_i/\lambda}}{d_i}\right]e^{j2\pi f_c t}\right\}, \tag{2.58}$$

where d_i denotes the total distance traveled by the ith ray, $\tau_i = d_i/c$, and G_i is the product of the transmit and receive antenna power gains corresponding to the ith ray. The first two terms in the summation of (2.58) correspond to the LoS and ground reflected path in the two-ray model (2.13), hence $R_0 = 1$ since there is no reflection of

[6] A rectilinear city is flat and has linear streets that intersect at 90° angles, as in midtown Manhattan.

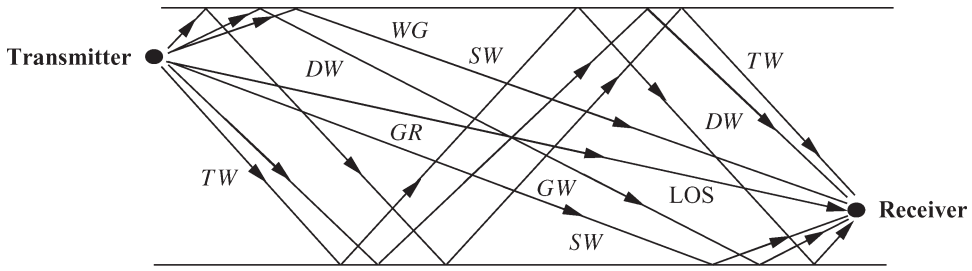


Figure 2.8: Overhead view of the ten-ray model. Includes the line-of-sight (LOS), ground-reflected (GR), two single-wall (SW) reflected, two double-wall (DW) reflected, two triple-wall (TW) reflected, wall-ground (WG) reflected, and ground-wall (GW) reflected rays

the LoS path. For the ith reflection path, $i > 0$, the coefficient R_i is either a single reflection coefficient given by (2.17) or, if the path corresponds to multiple reflections, the product of the reflection coefficients corresponding to each reflection. The dielectric constants used in (2.58) are approximately the same as the ground dielectric, so $\varepsilon_r = 15$ is used for all the calculations of R_i. The delay spread of the 10-ray and 6-ray models equals the difference between the delay of the LOS ray and that of the reflected ray with the largest delay: $\max_i \tau_i - \tau_0$.

If we again assume a narrowband model such that $u(t - \tau_0) \approx u(t - \tau_i)$ for all i, then the received power corresponding to (2.58) is

$$P_r = P_t \left[\frac{\lambda}{4\pi}\right]^2 \left|\sum_{i=0}^{9} \frac{R_i \sqrt{G_i} e^{-j\Delta\phi_i}}{d_i}\right|^2, \tag{2.59}$$

where $\Delta\phi_i = 2\pi(d_i - d_0)/\lambda$.

Power falloff with distance in both the ten-ray model (2.59) and urban empirical data [15, 32, 33] for transmit antennas both above and below the building skyline is typically proportional to d^{-2}, even at relatively large distances. Moreover, the falloff exponent is relatively insensitive to the transmitter height. This falloff with distance squared is due to the dominance of the multipath rays, which decay as d^{-2}, over the combination of the LOS and ground-reflected rays (two-ray model), which decays as d^{-4}. Other empirical studies [34, 35, 36] have obtained power falloff with distance proportional to $d^{-\gamma}$, where γ lies anywhere between 2 and 6.

2.8.2 Diffraction

The propagation model for the LOS and reflected paths was outlined in the previous section. Diffraction allows a radio signal to "bend around" an object in its path to the receiver, as shown in Figure 2.9 for a wedge-shaped object. This bending phenomenon can be explained using Huygen's principle, which states that all points on the signal wavefront can be considered as point sources for a secondary wavefront called a *wavelet*. When a signal is diffracted, these wavelets combine together to produce a wavefront in the new direction of propagation caused by the bending [29, Chapter 4.3]. Diffraction results from many phenomena, including hilly or irregular terrain, building and rooftop edges, or obstructions blocking the LOS path between the transmitter and receiver [2, 11, 16]. Diffraction can be accurately characterized using the geometrical theory of diffraction (GTD) [40], but the complexity of this approach has precluded its use in wireless channel modeling. Wedge diffraction simplifies the GTD by assuming the diffracting object is a wedge rather than a more general shape. This model has been used to characterize the mechanism by which signals are diffracted around street corners, which can result in signal attenuation exceeding 100 dB for some incident angles on the wedge [28, 39, 41, 42]. Although wedge diffraction simplifies the GTD, it still requires a numerical solution for the resulting attenuation [40, 43] and thus

is not commonly used. Diffraction is most commonly modeled by the *Fresnel knife-edge diffraction model* because of its simplicity. The geometry of this model is shown in Figure 2.9, where the diffracting object is assumed to be asymptotically thin, which is not generally the case for hills, rough terrain, or wedge diffractors. In particular, this model does not consider diffractor parameters such as polarization, conductivity, and surface roughness, which can lead to inaccuracies [41].

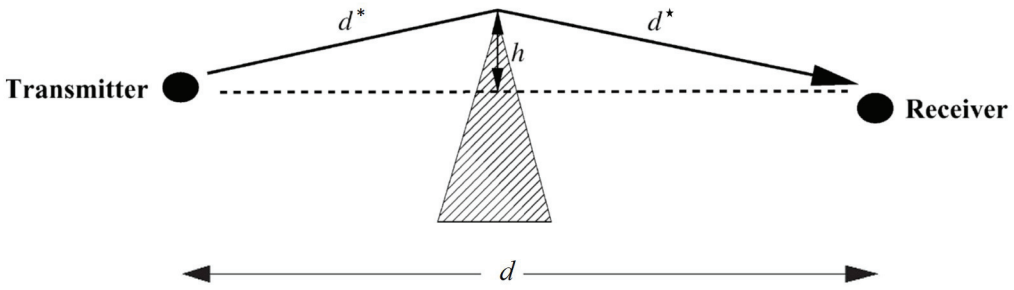


Figure 2.9: Knife-edge diffraction.

The diffracted signal of Figure 2.9 travels a distance $d^* + d^\star$, resulting in a phase shift of $\phi = 2\pi(d^* + d^\star)/\lambda$. For a LOS path of distance d, the geometry of Figure 2.9 indicates that, for h small relative to d^* and $d^\star$, the signal must travel an additional distance relative to the LOS path of approximately

$$\Delta d = d^* + d^\star - d \approx \frac{h^2(d^* + d^\star)}{2d^* d^\star}; \tag{2.60}$$

the corresponding phase shift relative to the LOS path is approximately

$$\Delta\phi = \frac{2\pi\Delta d}{\lambda} \approx \frac{\pi}{2}v^2, \tag{2.61}$$

where

$$v = h\sqrt{\frac{2(d^* + d^\star)}{\lambda d^* d^\star}} \tag{2.62}$$

is called the *Fresnel–Kirchhoff diffraction parameter*. The path loss of a signal at distance $d^* + d^\star$ experiencing knife-edge diffraction at distance d^* is generally a function of v. However, computing this path loss is fairly complex, requiring the use of Huygens's principle, Fresnel zones, and the complex Fresnel integral [2]. Moreover, the resulting diffraction loss cannot generally be found in closed form. Approximations for the additional path loss due to knife-edge diffraction relative to free-space path loss (in dB) as a function of v in (2.62) are given in [16, 44] as

$$L(v)\text{dB} = \begin{cases} 0, & v < -1, \\ 20\log_{10}[.5 - .62v], & -0.8 \le v < 0, \\ 20\log_{10}[.5e^{-.95v}], & 0 \le v < 1, \\ 20\log_{10}[.4 - \sqrt{.1184 - (.38 - .1v)^2}], & 1 \le v \le 2.4, \\ 20\log_{10}[.225/v], & v > 2.4. \end{cases} \tag{2.63}$$

The knife-edge diffraction received signal model is obtained by adding this loss to the free-space path loss of (2.11), yielding the following formula for the received diffracted signal:

$$r(t) = \mathrm{Re}\left\{ \frac{\lambda}{4\pi} \left[\frac{L(v)\sqrt{G_d}u(t-\tau)e^{-j2\pi(d^*+d^\star)/\lambda}}{d^* + d^\star} \right] e^{j2\pi f_c t} \right\}, \tag{2.64}$$

where $\sqrt{G_d}$ is the antenna gain and $\tau = (d^* + d^\star)/c$ is the delay associated with the diffracted ray.

In addition to diffracted rays, there may also be rays that are diffracted multiple times, or rays that are both reflected and diffracted. Models exist for including all possible permutations of reflection and diffraction [45]; however, the attenuation of the corresponding signal components is generally so large that these components are negligible relative to the noise. Diffraction models can also be specialized to a particular environment or frequency band. For example, a model for diffraction from rooftops and buildings in cellular systems was developed by Walfisch and Bertoni in [46]. This rooftop diffraction model was extended to multi-antenna base stations in [47], where it produced a "keyhole" phenomenon that reduced the effective spatial dimensions of the channel. Two-dimensional diffraction models for urban small cells, where signals propagate below rooftop levels, are developed in [48]. Application of knife-edge and more general diffraction models to millimeter wave propagation is discussed in [49, 50].

2.8.3 Scattering

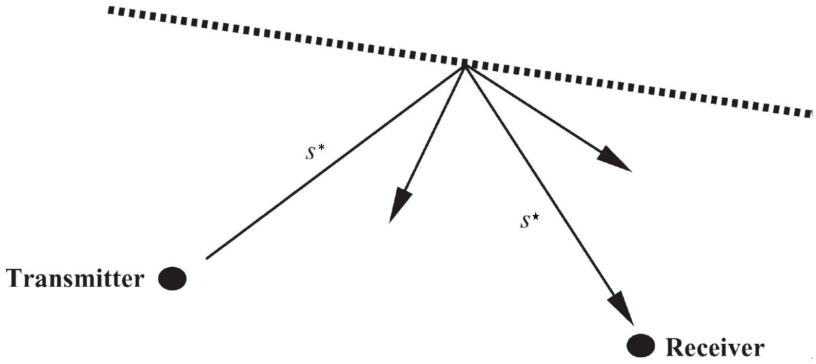


Figure 2.10: Scattering.

A scattered ray, shown in Figure 2.10 by the segments s_α and s_β, has a path loss proportional to the product of s_α and s_β. This multiplicative dependence is due to the additional spreading loss that the ray experiences after scattering. The received signal due to a scattered ray is given by the bistatic radar equation [51]:

$$r(t) = \mathrm{Re}\left\{ \frac{\lambda}{4\pi} \left[\frac{\sqrt{G_s}\sigma u(t-\tau)e^{-j2\pi(s^*+s^\star)/\lambda}}{\sqrt{4\pi}s^* s^\star} \right] e^{j2\pi f_c t} \right\}, \tag{2.65}$$

where $\tau = (s^* + s^\star)/c$ is the delay associated with the scattered ray; σ (in square meters) is the radar cross-section of the scattering object, which depends on the roughness, size, and shape of the scatterer; and G_s is the antenna power gain. The model assumes that the signal propagates from the transmitter to the scatterer based on free-space propagation and is then re-radiated by the scatterer with transmit power equal to σ times the received power at the scatterer. From (2.65), the path loss associated with scattering is

$$\begin{aligned} P_r \text{ dBm} = \quad & P_t \text{ dBm} + 10\log_{10}(G_s) + 20\log_{10}(\lambda) + 10\log_{10}(\sigma) \\ & -30\log(4\pi) - 20\log_{10}(s^*) - 20\log_{10}(s^\star). \end{aligned} \tag{2.66}$$

Empirical values of $10\log_{10}\sigma$ were determined in [52] for different buildings in several cities. Results from this study indicate that $10\log_{10}\sigma$ in dBm^2 ranges from -4.5 dBm^2 to 55.7 dBm^2, where dBm^2 denotes the dB value of the σ measurement with respect to one square meter.

2.8.4 Multipath Model with Reflection, Diffraction, and Scattering

The received signal is determined from the superposition of all the components due to the multiple rays. Thus, if we have a LOS ray, N_r reflected rays, N_d diffracted rays, and N_s scattered rays, the total received signal is

$$\begin{aligned} r_{\text{total}}(t) = \text{Re}\Bigg\{ \frac{\lambda}{4\pi} \Bigg[& \frac{\sqrt{G_0}u(t)e^{-j2\pi d_0/\lambda}}{d_0} + \sum_{i=1}^{N_r} \frac{R_i\sqrt{G_i}u(t-\tau_i)e^{-j2\pi d_i/\lambda}}{d_i} \\ & + \sum_{j=1}^{N_d} \frac{L_j(v)\sqrt{G_{d_j}}u(t-\tau_j)e^{-j2\pi(d_j^*+d_j^\star)/\lambda}}{d_j^* d_j^\star} \\ & + \sum_{k=1}^{N_s} \frac{\sqrt{G_{s_k}}\sigma_k u(t-\tau_k)e^{-j2\pi(s_k^*+s_k^\star)/\lambda}}{\sqrt{4\pi}s_k^* s_k^\star} \Bigg] e^{j2\pi f_c t} \Bigg\}, \end{aligned} \tag{2.67}$$

where τ_j and τ_k are, respectively, the time delays of the given diffracted and scattered rays, $(d_{\alpha_j}, d_{\beta_j})$ and $(s_{\alpha_k}, s_{\beta_k})$ are, respectively, the distance pair the signal travels before and after the object causing the diffraction or scattering, and the other parameters are as defined in the model (2.58) with the LoS and reflections only. The received power P_r of $r_{\text{total}}(t)$ and the corresponding path loss P_r/P_t are then obtained from (2.67). The delay spread of this model equals the difference between the delay of the LOS ray and that of the reflected, diffracted, or scattered ray that has the largest delay.

Any of these multipath components may have an additional attenuation factor if its propagation path is blocked by objects such as buildings outdoors or walls indoors. In this case, the attenuation factor of the obstructing object multiplies the component's path loss term in (2.67). This attenuation loss will vary widely, depending on the material and depth of the object [11, 53]. Models for random loss due to attenuation are described in Section 2.6.

2.8.5 Local Mean Received Power

The path loss computed from all ray-tracing models is associated with a fixed transmitter and receiver location. In addition, ray tracing can be used to compute the *local mean received power* $\overline{P}_r$ in the vicinity of a given receiver location by adding the squared magnitude of all the received rays. This has the effect of averaging out local spatial variations due to phase changes around the given location. Local mean received power is a good indicator of link quality and is often used in cellular system functions like power control and handoff [86].

2.9 Directional Antennas, Antenna Arrays, and Intelligent Surfaces

The two-ray model introduced the notion of directional antennas, which focus energy from the transmitter or receiver in particularly directions. This directionality can increase the signaling range, reduce delay spread and

flat fading, and suppress interference between users. In particular, interference typically arrives at the receiver from different directions. Thus, directional receive antennas can null or attenuate interference arriving from given directions, thereby increasing the signal-to-interference power ratio. The reflected multipath components of the transmitted signal also arrive at the receiver from different directions and thus can be attenuated by directional receive antennas, thereby reducing ISI and flat fading.

The antenna gain pattern of a directional antenna can be static or dynamic. Common static designs include parabolic dish antennas, Yagi antennas, sectorized antennas, and static phased arrays. Dish and Yagi antennas are relatively large and are used primarily to amplify satellite signals as well as those in the HF, UHF, and VHF frequency bands. Cellular and Wi-Fi systems tend to use sectorized antennas or phased arrays whose gain patterns, along with that of an omnidirectional antenna gain, are shown in Figure 2.11.

Sectorized antennas are designed to provide high gain across a range of signal arrival angles. These antennas are generally too large to fit on a mobile device. Hence sectorization is commonly used at a cellular system base stations to cut down on interference: assuming different sectors are assigned the same frequency band or timeslot, then with perfect sectorization only users within a sector interfere with each other, thereby reducing the average interference by a factor equal to the number of sectors. For example, Figure 2.11 shows a sectorized antenna with a 120° beamwidth. A base station could divide its 360° angular range into three sectors to be covered by three 120° sectorized antennas, in which case the interference in each sector is reduced by a factor of 3 relative to an omnidirectional base station antenna. The price paid for reduced interference in cellular systems via sectorization is the need for handoff between sectors.

Phased antenna arrays are the most common method to create directional antennas in Wi-Fi and cellular systems since they are adaptable and, at high enough frequencies where antennas sizes are small, a phased array can be incorporated into a mobile device. Phased array techniques work by adapting the phase of each antenna element in the array to steer energy in a particular direction. This technique is also known as beamsteering, since the phased array beams energy in a given direction, The larger the array, the more precisely energy can be steered in a particular direction. The phases of each antenna element in a phased array is based on a known angular location of the desired and interfering signals in order to provide high or low gains in the appropriate directions. Tracking of user locations can be a significant impediment in highly mobile systems, which is why cellular base stations use sectorized instead of directional antennas.

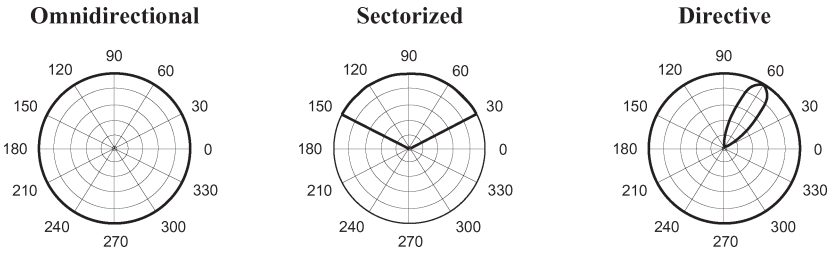


Figure 2.11: Antenna gains for omnidirectional, sectorized, and directive antennas.

Ray tracing models for systems with phased arrays at the transmitter and/or receiver utilize superposition. In particular, the single-antenna ray tracing technique described in Sections 2.8.1- 2.8.4 is calculated for each transmit-receive antenna pair and the resulting signals summed together. Equivalently, the antenna gain pattern for the transmit and receive antenna array can be determined and the transmit and receive directional gain applied to each ray propagating between the transmitter and receiver. This equivalent method was used in the two-ray model described in Section 2.4 above.

When both the transmitter and receiver utilize phased arrays, the system is called a multiple-input multiple output (MIMO) system. As described in detail in Chapter 10, by utilizing antenna arrays at both the transmitter and

receiver, MIMO system can provide directional antenna gain at both the transmitter and receiver. Alternatively, MIMO systems can used their transmit and receive antenna arrays to provide diversity gain against fading, create parallel channels to increase data rate, or to null out interferers. While the application of ray tracing to MIMO systems is straightforward, the computational complexity of this approach grows linearly with the product of the number of transmit and receive antennas. As a result, ray tracing approximations to lower this complexity have been developed, including clustering of rays [54] and preprocessing environmental data [58]. A summary of ray tracing methods for MIMO systems can be found in [55].

MIMO systems and directional antennas provide antenna gain in a particular direction through beamsteering at the transmitter and receiver. Intelligent surfaces take a different approach; they use antenna arrays deployed in a signal's propagation environment to steer signal reflections in the direction of one or more receivers. Most intelligent surfaces are reconfigurable, using a large number of programmable passive elements whose phases are dynamically tuned to create a directed beam of the signal reflected off its surface, as shown in Fig. 2.12. These *Reconfigurable Intelligent Surfaces (RIS)* are most effective in propagation environments where the LOS signal may be blocked, since the reflected signal has the same characteristics as the LOS signal but coming from an unblocked direction. When the LOS signal is not blocked, the RIS can tune the phase of the reflected signal to combine coherently with that signal to increase its power. A RIS can also reduce interference at a given receiver by tuning the reflected interference signal such that it non-coherently combines with its LOS signal. RISs can be installed indoors or outdoors on any large flat surface, e.g. on the outer wall of a building or on an inside ceiling or wall. Because the reflecting elements are passive phase shifters, i.e. they do not have active RF chains, they are both low-cost and low-power. Challenges to incorporate RISs into indoor and outdoor wireless systems include where and how to deploy the surfaces, channel estimation and the impact of estimation error in a dynamically changing environment, as well as the design of low-complexity distributed and centralized algorithms for beamsteering. More discussion of RISs and their design challenges can be found in [56, 57].

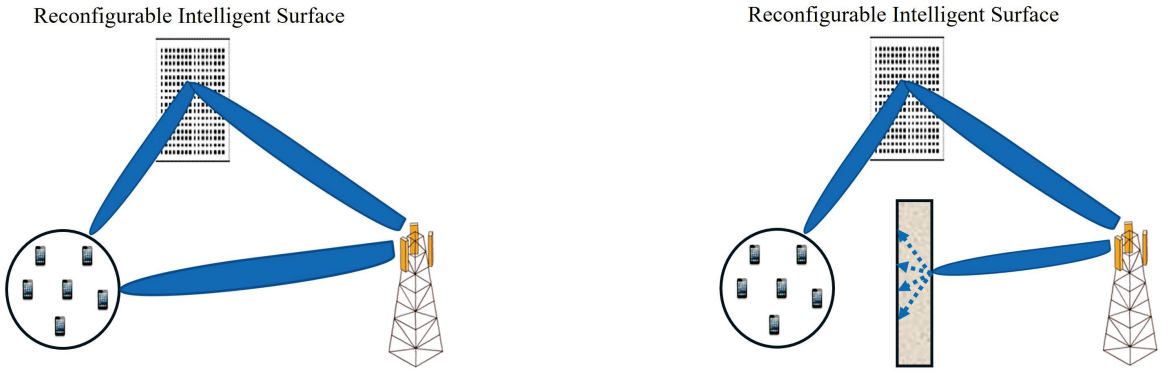


Figure 2.12a: RIS with unblocked LOS Signal

Figure 2.12b: RIS with blocked LOS Signal

Figure 2.12: Beamsteering of a reflected signal using a reconfigurable intelligent surface (RIS).

2.10 Measurement-Based Propagation Models

Most mobile communication systems operate in complex propagation environments that cannot be accurately modeled by free-space path loss or ray tracing. A number of analytical path loss models have been developed over the years by researchers as well as standards bodies to model path loss in typical wireless environments including urban macrocells and small cells, suburban areas, rural areas, and inside buildings with different characteristics

[11, 59, 61, 77, 17]. These models are generally based on large empirical measurement campaigns that can range over a variety of distances, frequency ranges, geographical regions for outdoor models, and building types for indoor models. These analytical models have the highest accuracy when they are applied to propagation conditions similar to those under which the empirical measurements that the models are based on were made. In this section we describe the most common analytical path loss models based on empirical measurements for both indoor and outdoor systems.

Analytical models characterize P_r/P_t as a function of distance, so path loss is well-defined. In contrast, empirical measurements of P_r/P_t as a function of distance include the effects of path loss, shadowing, multipath, and other site-specific factors that affect propagation. In order to remove multipath effects, empirical measurements for path loss typically average their received power measurements and the corresponding path loss at a given distance over several wavelengths. This average path loss is called the *local mean attenuation* (LMA) at distance d, and it generally decreases with d owing to free-space path loss and signal obstructions. The LMA in a given environment, such as a city, depends on the specific location of the transmitter and receiver corresponding to the LMA measurement. To characterize LMA more generally, measurements are typically taken throughout the environment and possibly in multiple environments with similar characteristics. Thus, the *empirical path loss* $P_L(d)$ for a given environment (a city, suburban area, or office building) is defined as the average of the LMA measurements at distance d averaged over all available measurements in the given environment. For example, empirical path loss for a generic downtown area with a rectangular street grid might be obtained by averaging LMA measurements in New York City, downtown San Francisco, and downtown Chicago. The empirical path loss models given in this section are all obtained from average LMA measurements. Empirical path loss models can also be developed or refined using measurement-driven learning applied to continuous LMA data collection [78].

2.10.1 Okumura Model

One of the most well-known models for signal prediction in large urban macrocells is the Okumura model [87]. This model is applicable over distances of 1–100 km and frequency ranges of 150–1500 MHz. Okumura used extensive measurements of base station-to-mobile signal attenuation throughout Tokyo to develop a set of curves giving median attenuation relative to free space of signal propagation in irregular terrain. The base station heights for these measurements were 30–100 m, a range whose upper end is higher than typical base stations today. The empirical path loss formula of Okumura at distance d parameterized by the carrier frequency f_c is given by

$$P_L(d)\ \text{dB} = L(f_c, d) + A_\mu(f_c, d) - G(h_t) - G(h_r) - G_{\text{AREA}}, \tag{2.68}$$

where $L(f_c, d)$ is free-space path loss at distance d and carrier frequency f_c, $A_\mu(f_c, d)$ is the median attenuation in addition to free-space path loss across all environments, $G(h_t)$ is the base station antenna height gain factor, $G(h_r)$ is the mobile antenna height gain factor, and G_{AREA} is the gain due to the type of environment. The values of $A_\mu(f_c, d)$ and G_{AREA} are obtained from Okumura's empirical plots [87, 11]. Okumura derived empirical formulas for $G(h_t)$ and $G(h_r)$ as follows:

$$G(h_t) = 20\ \log_{10}(h_t/200), \quad 30\ \text{m} < h_t < 1000\ \text{m}; \tag{2.69}$$

$$G(h_r) = \begin{cases} 10 \log_{10}(h_r/3), & h_r \leq 3\ \text{m}, \\ 20 \log_{10}(h_r/3), & 3\ \text{m} < h_r < 10\ \text{m}. \end{cases} \tag{2.70}$$

Correction factors related to terrain are also developed in [87] that improve the model's accuracy. Okumura's model has a 10–14 dB empirical standard deviation between the path loss predicted by the model and the path loss associated with one of the measurements used to develop the model. The expected error in using Okumura's model for environments not based on these measurements is generally higher.

2.10.2 Hata Model

The Hata model [88] is an empirical formulation of the graphical path loss data provided by Okumura and is valid over roughly the same range of frequencies, 150–1500 MHz. This empirical model simplifies calculation of path loss because it is a closed-form formula and is not based on empirical curves for the different parameters. The standard formula for empirical path loss in urban areas under the Hata model is

$$\begin{aligned} P_{L,\text{urban}}(d)\ \text{dB} = 69.55 + 26.16\log_{10}(f_c) - 13.82\log_{10}(h_t) - a(h_r) \\ + (44.9 - 6.55\log_{10}(h_t))\log_{10}(d). \end{aligned} \tag{2.71}$$

The parameters in this model are the same as under the Okumura model, and $a(h_r)$ is a correction factor for the mobile antenna height based on the size of the coverage area [88, 11]. For small to medium-sized cities, this factor is given by

$$a(h_r) = (1.1\log_{10}(f_c) - .7)h_r - (1.56\log_{10}(f_c) - .8)\ \text{dB}, \tag{2.72}$$

For larger cities this factor is given by

$$a(h_r) = 3.2(\log_{10}(11.75h_r))^2 - 7.63\ \text{dB} \tag{2.73}$$

for frequencies above 400 MHz and by

$$a(h_r) = 8.29(\log_{10}(1.54h_r))^2 - 3.69\ \text{dB} \tag{2.74}$$

for frequencies below 200 MHz.

Corrections to the urban model are made for suburban and rural propagation, so that these models are, respectively,

$$P_{L,\text{suburban}}(d)\ \text{dB} = P_{L,\text{urban}}(d)\ \text{dB} - 2[\log_{10}(f_c/28)]^2 - 5.4 \tag{2.75}$$

and

$$P_{L,\text{rural}}(d)\ \text{dB} = P_{L,\text{urban}}(d)\ \text{dB} - 4.78[\log_{10}(f_c)]^2 + 18.33\log_{10}(f_c) - K, \tag{2.76}$$

where K is 40.94 for an open area. Unlike the Okumura model, the Hata model does not provide for any path-specific correction factors. The Hata model well approximates the Okumura model for distances $d > 1$ km. Hence it is a good model for early-generation cellular systems operating in cities, but it does not model propagation well in cellular systems with smaller cell sizes, multiple antennas, or higher frequencies. Indoor environments are also not captured by the Hata model. As cellular systems evolved beyond their early deployments, research into more complex propagation models was undertaken by multiple organizations, as described in the next section.

2.10.3 Cellular System Channel Models

Following the success of early cellular systems, funding agencies and standards bodies worldwide launched projects to develop standardized cellular system channel models for the environments, carrier frequencies, cell sizes, and multiple antenna characteristics of these evolving systems. These standardized channel models can be used to evaluate different candidate cellular technologies in a uniform manner, and also provide guidelines for expected system performance in typical environments and operating conditions. Cellular system channel models span frequency ranges from .5 GHz to 100 GHz, indoor and outdoor environments with both LOS and non-LOS propagation, small and large cells sizes, and multiple antenna transmitters and receiver of different antenna heights, elevation angles, and polarizations. Models for all of these scenarios have been developed by 3GPP and the ITU

[59, 60], whereas other standardized models are applicable to a subset of these settings. To illustrate some basic properties of these standardized models, in this section we provide the LOS path loss formulas for a few standardized models, which follow either the single-slope or dual-slope formulas ((2.25) and (2.31), respectively) with model-specific constants. Extensions of these models to more complex environments are also discussed. More details on the unique characteristics of millimeter wave propagation are provided in Section 2.10.5.

An early project to standardize channel models for cellular systems was undertaken by the European Cooperation in Science and Technology (COST). The COST 207 program, launched in 1984, brought together industry, government, and academia under a common umbrella to develop common propagation models for the emerging 2G cellular systems. The subsequent COST 231 project was focused on extending earlier channel models to carrier frequencies up to 2 GHz. The resulting model, referred to as the COST 231 extension to the Hata model [89], has the following path loss expression:

$$\begin{aligned} P_{L,\,\text{urban}}(d)\ \text{dB} = {} & 46.3 + 33.9\log_{10}(f_c) - 13.82\log_{10}(h_t) - a(h_r) \\ & + (44.9 - 6.55\log_{10}(h_t))\log_{10}(d) + C_M, \end{aligned} \tag{2.77}$$

where $a(h_r)$ is the same correction factor as in the Hata model and $C_M = 0$ dB for medium-sized cities and suburbs while $C_M = 3$ dB for metropolitan areas. The COST 231 extension to the Hata model is restricted to the following range of parameters: 1.5 GHz $< f_c <$ 2 GHz, $30\,\text{m} < h_t < 200\,\text{m}$, $1\,\text{m} < h_r < 10\,\text{m}$, and 1 km $< d <$ 20 km. Given the Okumura measurements on which it is based, this model is most accurate for heterogeneous macrocell-only cellular system architectures with single-antenna base stations and terminals. The COST 231 model was further extended to incorporate new features and environments of the evolving cellular systems and standards, including diffraction and non-LOS propagation (COST-231-WI), multiple base station antennas (COST 259) [91], and multiple antennas at both the base stations and the mobile terminals (COST 273, COST 2100) [92, 93].

The 3GPP standards body maintains a family of channel models characterizing signal propagation for the different environments and frequencies in which their cellular systems operate [59]. These include path loss models for both LOS and non-LOS propagation in outdoor macrocells and small cells as well as for indoor office buildings. The 3GPP outdoor propagation models follow the dual-slope path loss model (2.31); they have a distance-dependent path loss exponent that takes on two different values, γ_1 and γ_2, for distances that are, respectively, smaller or greater than the breakpoint distance. The breakpoint distance is set to the critical distance $d_c = 4h_th_r/\lambda$ given by (2.23) in the two-ray model based on the effective transmit (h_t) and receive (h_r) antenna heights[7]. In particular, the 3GPP path loss for LOS propagation in outdoor macrocells and small cells is given by

$$P_L(d)\,\text{dB} = \begin{cases} 20\log_{10}(f_c) + \kappa + 10\gamma_1\log_{10}(d), & d \le d_c, \\ 20\log_{10}(f_c) + \kappa + 10\gamma_2\log_{10}(d) + \eta\log_{10}\left(d_c^2 + (h_t - h_r)^2\right), & d > d_c, \end{cases} \tag{2.78}$$

for d in meters and f_c in GHz. The 3GPP indoor office model is governed by the first terms of (2.78) for all distances (i.e., there is no breakpoint distance in the indoor model due to the shorter distances involved). At distances below d_c this model has three separate terms: the constant κ that captures environment and antenna parameters independent of distance and frequency, the term $20\log_{10} f_c$ that captures frequency-dependent path loss, and the term $10\gamma_i\log_{10}(d)$ that captures distance-dependent path loss. The frequency-dependent path loss $20\log_{10} f_c$ is the same as in free space (2.10), and this term along with κ is captured by the constant K in the single-slope model (2.25). At distances above d_c the macrocell and small cell models have a fourth term that captures path loss dependence on antenna heights. The values of κ, γ_i, and η for the macrocell, small cell, and indoor office models are given in Table 2.4.

[7] In the 3GPP model the effective antennas heights are the actual heights reduced by 1m.

Table 2.4: Parameters for 3GPP LOS Path Loss Models

Parameter	Macrocell	Small Cell	Indoor Office
κ	28	32.4	32.4
γ	$\gamma_1 = 22, \gamma_2 = 40$	$\gamma_1 = 21, \gamma_2 = 40$	$\gamma_1 = 17.3$
η	-9	-9.5	NA

The 3GPP family of channel models also include non-LOS propagation models that take path loss to be the maximum of the LOS path loss model and a different non-LOS path loss model. The non-LOS path loss model is of the same form as (2.78) but with different parameters and a fourth term that depends only on the receiver height h_r. There is also a 3GPP rural macrocell model that includes additional parameters of the environment such as building heights and street widths. The 3GPP Spatial Channel Model (SCM) incorporates MIMO parameters. Details can be found in [59]. The ITU channel models are similar to the 3GPP channel family and cover the same frequency range of .5-100 GHz [60].

The European WINNER (Wireless World Initiative New Radio) I and II projects created a similar family of channel models as 3GPP for cellular system propagation under different conditions and environments [61]. The WINNER path loss model is similar to the single-slope model (2.25) with the constant K broken into 3 separate terms. Specifically, the path loss in the WINNER propagation model $10\log_{10}(P_t/P_r)$ is given by

$$P_L\text{dBm} = 10\gamma\log_{10}(d) + \kappa + \beta\log_{10}(.2f_c) + X \tag{2.79}$$

for d in meters and f_c in GHz, where γ is the path loss exponent, κ is the path loss at the antenna near-field reference distance, typically a few signal wavelengths [12], which is independent of frequency, $\beta\log_{10}(.2f_c)$ is the frequency-dependence term, and X is an optional environment-specific term that may, for example, capture a fixed attenuation at all distances. The Winner II model includes MIMO channel parameters.

The COST, 3GPP, ITU, and WINNER families of channel models are widely used by researchers and practitioners for modeling path loss in cellular systems at frequencies below 6 GHz, with the 3GPP and ITU models covering frequencies up to 100 GHz. These models also include delay spread parameters ranging from tens of nanoseconds for indoor environments to hundreds of nanoseconds for outdoor environments. Alternative empirical channel models specifically for MIMO systems include the WIMAX 802.16 and Stanford University Interim (SUI) channel models [79]. As cellular system architectures have evolved to include large and small cells, device-to-device communications, base stations mounted on UAVs and drones, large antenna arrays, millimeter wave frequencies, and very high speed mobile devices, there have been extensive measurement campaigns to characterize propagation under these new configurations and conditions [80, 81, 82]. In addition, learning algorithms that continually update channel models based on collected data have also been developed [83, 84, 85]. Clearly measurement-based models for cellular systems will continue to evolve as new system architectures, frequency bands, device requirements, and channel measurement tools are introduced.

2.10.4 Wi-Fi Channel Models

Prior to 802.11n, performance evaluation of 802.11 WLAN systems was done using channel models for 2.4 GHz (802.11g) and 5 GHz (802.11a) that had been developed by individual researchers or companies. There was no general consensus on which models to use for performance evaluation of these earlier systems. To address this shortcoming, the 802.11n task group (TG) was formed to develop a standard set of models, called the TGn models, to evaluate 802.11n systems. As per the specifications of the 802.11n standard, these models include both 20 MHz and 40 MHz channels with up to 4 antennas at the transmitter and receiver. Six different TGn channel models (models A-F) comprise the TGn family to characterize propagation in different size homes and offices as well

as outdoors. All six follow the dual-slope path loss model (2.31) with a path loss exponent $\gamma_1 = 2$ prior to the breakpoint distance $d_{\rm BP}$ and $\gamma_2 = 3.5$ above this distance. The breakpoint distance $d_{\rm BP} = 5$ m in models A, B, and C, which correspond to LOS only, small home, and large home, respectively. This distance ranges from 10-30 m in models D-F (corresponding to small office, large office, and outdoors, respectively). The delay spread of the multipath, normalized with respect to the LOS path, becomes the difference between the delay of the multipath component with the largest delay and the LOS path. In these models this normalized delay spread ranges from 0 ns for the A model to 150 ns for the F model. Multipath reflections are assumed to arrive in clusters based on the indoor propagation model developed in [62]. Antenna correlation is modeled using the Kronecker model described in [63]. Shadow fading follows the log-normal model with standard deviation $\sigma_{\psi_{\rm dB}} = 3$ for distances $d < d_{\rm BP}$. For $d > d_{\rm BP}$, $\sigma_{\psi_{\rm dB}} = 4$ for Models A and B, $\sigma_{\psi_{\rm dB}} = 5$ for Models C and D, and $\sigma_{\psi_{\rm dB}} = 6$ for Models E and F. More details on the TGn models can be found in [64].

The TGac family of channels extends the TGn model to support performance evaluation of 802.11ac systems. This model increases the number of transmit and receive antennas using the same correlation model as TGn, increases channel bandwidth up to 160 MHz, and modifies the multipath clustering models to account for simultaneous transmission to multiple users (multiuser MIMO). The TGax family of channel models provides additional environmental scenarios for dense deployments both indoors and outdoors. Details of these extensions can be found in [65, 66].

2.10.5 Millimeter Wave Models

The millimeter (mmW) frequency band, spanning the 30-300 GHz range, was not used in the first few generations of Wi-Fi and cellular systems due to its high path loss as well as the high cost and poor performance of mmW hardware. However, as demand for increased data rates in wireless systems continued unabated with each new generation, utilizing the large amount of available spectrum in this band grew increasingly appealing. Moreover, in 2005 the FCC in the US created more flexible rules for spectrum allocation in the lower part of this band to increase its use [67]. Following this ruling, a few commercial mmW communication systems were developed for in-home use, but the technology was not yet competitive with Wi-Fi. Interest in utilizing mmW wave spectrum for cellular communications also started to grow, leading to outdoor channel measurement campaigns as well as cellular system designs to overcome the propagation challenges at these high frequencies. This section provides an overview of mmW propagation characteristics.

Since path loss is inversely proportional to frequency squared in the free space path loss formula (2.6), this loss is much higher at mmW frequencies than at lower frequencies. However, this large free-space path loss can be compensated by high gains associated with directional or multiple antenna technologies. In addition to free-space path loss, signal propagation in the mmW band experiences significant attenuation at specific carrier frequencies due to interactions with oxygen, water, and other molecular components in the air. Attenuation due to shadowing from objects and people is more severe at these frequencies and shadowing can also cause scattering of directed beams. As a result, there is typically only a small number of multipath components in mmWave channels whose received power is above the noise floor.

The resonant frequency of oxygen (O_2) molecules leads to heavy absorption, up to 15 dB/km, for signals transmitted through the air at carrier frequencies around 60 GHz. There is also significant absorption due to the resonant frequencies of O_2 or water vapor (H_2O) at 119, 183, and 325 GHz, with a small peak at 24 GHz. The combined effects of free-space path loss and the O_2 and H_2O absorption in non-rainy conditions is illustrated in Fig. 2.13 from [70]. It can be seen from this figure that atmospheric absorption is somewhat larger at sea level than at higher altitudes due to increased water vapor concentration.

Millimeter wave propagation is also affected by rain absorption. In particular, rain causes scattering since the wavelength of a mmW signal is roughly the same size as a raindrop. The higher the rain rate, the more scattering occurs and hence the higher the absorption. A model for attenuation at different rain rates for mmW propagation

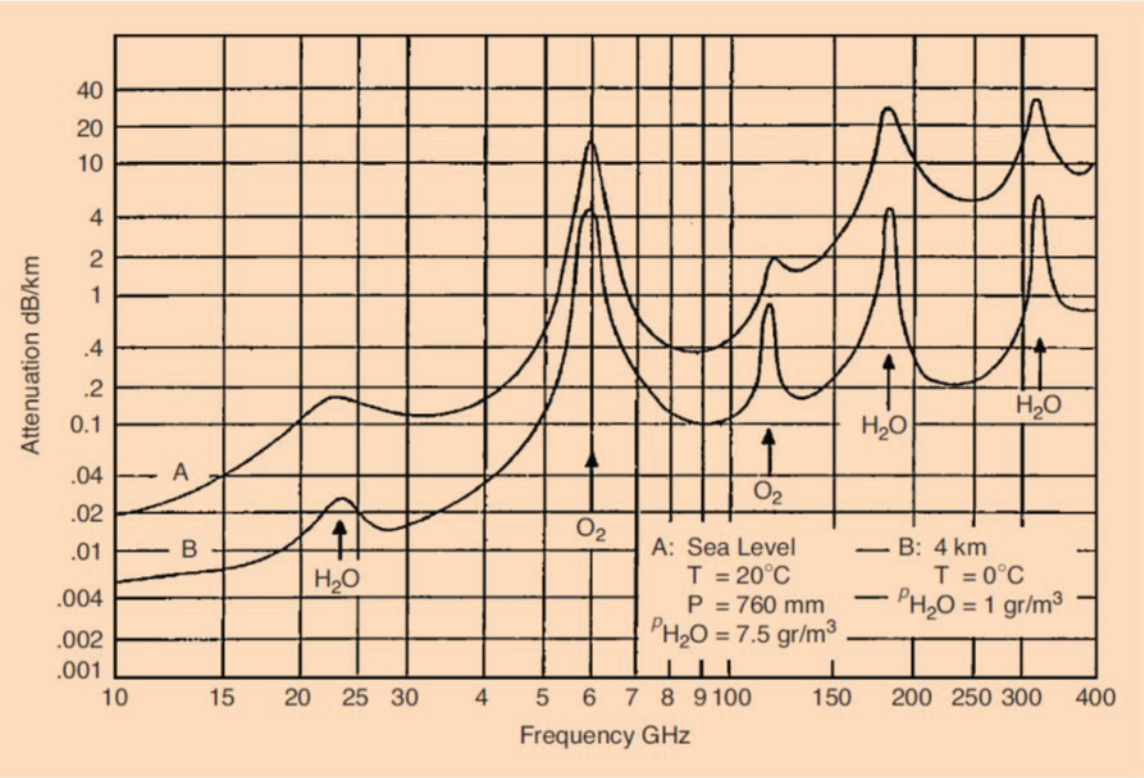


Figure 2.13: Path Loss, O2, and H2O Absorption in Millimeter Wave Propagation [70].

was developed by the ITU [71], as illustrated in Fig. 2.14 from [68]. A simple approximation, developed in [72], gives the rain attenuation as $.95R^{.77}$ for R the rain rate in mm/hr.

While atmospheric absorption and rain attenuation entail very large path loss over large distances, most mmWave systems are designed to operate over relatively short distances, less than 500 m. The losses illustrated in Figs. 2.13-2.14 for these short distances can be compensated through a combination of increased transmit power, antenna techniques, modulation choice, and channel coding [69].

As interest in mmW frequency utilization for cellular and Wi-Fi grew, extensive measurement campaigns were undertaken in different indoor and outdoor environments [73], with analytical channel models developed to approximate these empirical measurements. The most widely known of these models, summarized in [74], are based on work done by four different organizations: the 3GPP standards body, the 5G Channel Model (5GCM) ad hoc group, the Mobile and wireless communications Enablers for the Twenty-twenty (2020) Information Society (METIS) research project, and the Millimeter-Wave Based Mobile Radio Access Network for 5G Integrated Communications (mmMAGIC) research project. The models for each group capture propagation for four different categories of environments: outdoor macrocell, outdoor small cell, home, and commercial building. In addition to models for path loss and shadowing, these mmW propagation models add a new component called *LOS probability*, which captures the probability that the LOS signal path between the transmitter and receiver is not blocked or heavily attenuated. While this probability is different for the four environment categories within each family of models, in all cases it depends on the transmit-receive distance and does not depend on the carrier frequency [75]. Millimeter wave path loss in all these models is assumed to follow the dual-slope model (2.31), where the breakpoint distance is different for each of the four environments. The shadowing standard deviation for each of

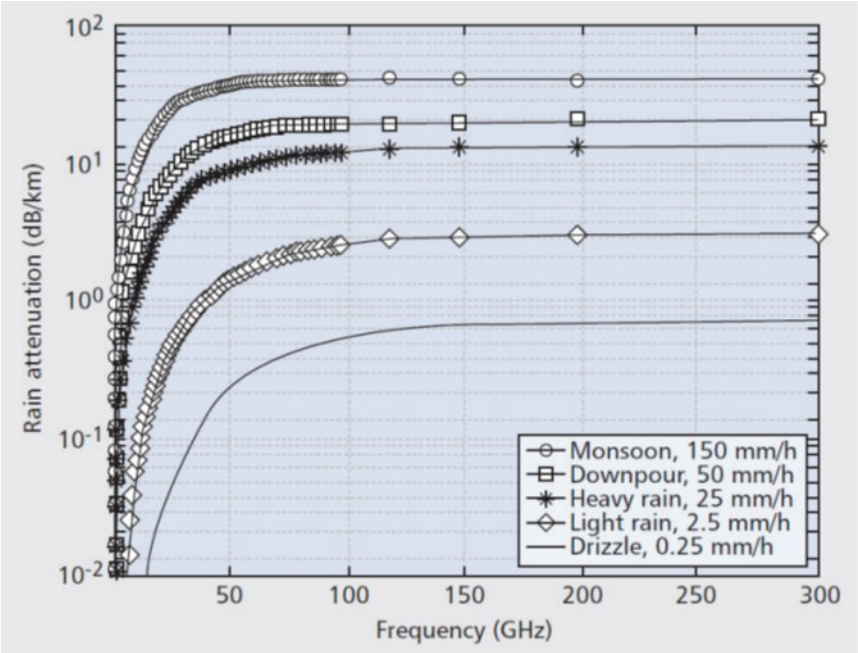


Figure 2.14: Rain Attenuation in Millimeter Wave Propagation [68].

the four environments is a fixed constant. As expected, shadowing in these models is larger for non-LOS compared to LOS environments, and ranges from about 1-8 dB. A summary of the differences between mmW and microwave channel models can be found in [75]. The sensitivity of performance analysis to the particular channel model adopted is demonstrated in [82]. The terahertz (THz) band (.3-3 THz) offers even more spectrum than the mmW band, along with more challenging propagation characteristics [76].

2.10.6 Indoor Attenuation Models

Indoor environments differ widely in the materials used for walls and floors, the layout of rooms, hallways, windows, and open areas, the location and material in obstructing objects, the size of each room, and the number of floors. All of these factors have a significant impact on path loss in an indoor environment. While the TGn, TGac, and TGax Wi-Fi models capture propagation in several generic indoor environments, the nature of these generic models make them inaccurate in characterizing propagation for any particular indoor environment. In this section we describe indoor models that can be tailored to such environments.

Indoor path loss models must accurately capture the effects of attenuation across floors due to partitions as well as between floors. Measurements across a wide range of building characteristics and signal frequencies indicate that the attenuation per floor is greatest for the first floor that is passed through and decreases with each subsequent floor. Specifically, measurements in [18, 97, 19, 20] indicate that, at 900 MHz, the attenuation when transmitter and receiver are separated by a single floor ranges from 10–20 dB, while subsequent attenuation is 6–10 dB per floor for the next three floors and then a few decibels per floor for more than four floors. At higher frequencies the attenuation loss per floor is typically larger [97, 98]. The attenuation per floor is thought to decrease as the number of attenuating floors increases because of the scattering up the side of the building and reflections from adjacent buildings. Partition materials and dielectric properties vary widely and thus so do partition losses. Measurements for the partition loss at different frequencies for different partition types can be found in [11, 18, 99, 100, 23], and

Table 2.5: Typical partition losses at $f_c \approx 1$ GHz

Partition type	Partition loss (dB)
Cloth partition	1.4
Double plasterboard wall	3.4
Foil insulation	3.9
Concrete wall	13
Aluminum siding	20.4
All metal	26

Table 2.6: ITU penetration loss model

Material	Penetration loss (dB)
Standard multi-pane glass	$2 + 0.2f_c$
Infrared reflective glass	$23 + .3f_c$
Concrete	$5 + 4f_c$
Wood	$4.85 + 0.12f_c$

Table 2.5 indicates a few examples of partition losses measured at around 1 GHz from this data. The partition loss obtained by different researchers for the same partition type at the same frequency often varies widely, so it is difficult to make generalizations about partition loss from a specific data set.

The experimental data for floor and partition loss can be added to an analytical or empirical dB path loss model $P_L(d)$ as

$$P_r \,\mathrm{dBm} = P_t \,\mathrm{dBm} - P_L(d) - \sum_{i=1}^{N_f} \mathrm{FAF}_i - \sum_{i=1}^{N_p} \mathrm{PAF}_i, \tag{2.80}$$

where FAF_i represents the floor attenuation factor for the ith floor traversed by the signal and PAF_i represents the partition attenuation factor associated with the ith partition traversed by the signal. The number of floors and partitions traversed by the signal are N_f and N_p, respectively.

Another important factor for indoor systems whose transmitter is located outside the building is outdoor-to-indoor penetration loss. Measurements indicate that this penetration loss is a function of frequency, height, and the exterior wall or window material. For carrier frequencies from .9-3 GHz penetration loss measurements range from 8 dB to 20 dB [2, 21, 103, 104]. Penetration loss for mmWave signals generally increases with frequency. Hence, the 3GPP, ITU, 5GCM, and mmMAGIC propagation models include a component for outdoor-to-indoor penetration loss as a function of both frequency and the exterior wall or window material [59, 60, 73]. The ITU penetration loss model as a function of frequency f_c and for different wall and window materials is given in Table 2.6. In general penetration loss decreases by about 1.4 dB per floor at floors above the ground floor. This decrease is due to reduced clutter at higher floors and the higher likelihood of a LOS path.

Chapter 2 Problems

1. Under the free-space path loss model, find the transmit power required to obtain a received power of 1 dBm for a wireless system with isotropic antennas ($G_t = G_r = 1$) and a carrier frequency $f_c = 5$ GHz, assuming a distance $d = 10\,\text{m}$. Repeat for $d = 100\,\text{m}$.

2. For the two-ray model with transmitter–receiver separation $d = 100\,\text{m}$, $h_t = 10\,\text{m}$, and $h_r = 2\,\text{m}$, find the delay spread between the two signals.

3. For the two-ray model, show how a Taylor series approximation applied to (2.15) results in the approximation

$$\Delta\phi = \frac{2\pi(d_1 - d_0)}{\lambda} \approx \frac{4\pi h_t h_r}{\lambda d}.$$

4. For the two-ray model, derive an approximate expression for the distance values below the critical distance d_c at which signal nulls occur.

5. Find the critical distance d_c under the two-ray model for a large macrocell in a suburban area with the base station mounted on a tower or building ($h_t = 20\,\text{m}$), the receivers at height $h_r = 3\,\text{m}$, and $f_c = 2$ GHz. Is this a good size for cell radius in a suburban macrocell? Why or why not?

6. Suppose that, instead of a ground reflection, a two-ray model consists of a LOS component and a signal reflected off a building to the left (or right) of the LOS path. Where must the building be located relative to the transmitter and receiver for this model to be the same as the two-ray model with a LOS component and ground reflection?

7. Consider a two-ray channel with impulse response $h(t) = \alpha_1\delta(t) + \alpha_2\delta(t - .022\,\mu\text{s})$, so for input signal $s(t)$ to the channel, the output is $r(t) = h(t) * s(t)$. Find the distance separating the transmitter and receiver, as well as α_1 and α_2, assuming free-space path loss on each path with a reflection coefficient of -1. Assume the transmitter and receiver are located 8 m above the ground and that the carrier frequency is 900 MHz.

8. Directional antennas are a powerful tool to reduce the effects of multipath as well as interference. In particular, directional antennas along the LOS path for the two-ray model can reduce the attenuation effect of ground wave cancellation, as will be illustrated in this problem. Assume the reference distance $d_r = 1$ m. Plot the dB power ($10\log_{10} P_r$) versus log distance ($\log_{10} d$) for the two-ray model with parameters $f_c = 900\,\text{MHz}$, $R = -1$, $h_t = 50\,\text{m}$, $h_r = 2\,\text{m}$, $G_0 = 1$, and the following values for $G_1 : G_1 = 1, .316, .1,$ and .01 (i.e., $G_1 = 0,\ -5,\ -10,$ and -20 dB, respectively). Each of the four plots should range in distance from $d = 1$ m to $d = 100$ km. Also calculate and mark the critical distance $d_c = 4h_t h_r/\lambda$ on each plot, and normalize the plots to start at approximately 0 dB. Finally, show the piecewise linear model with flat power falloff up to distance h_t, falloff $10\log_{10}(d^{-2})$ for $h_t < d < d_c$, and falloff $10\log_{10}(d^{-4})$ for $d \geq d_c$. (On the power loss versus log distance plot, the piecewise linear curve becomes a set of three straight lines of slope $0, 2,$ and 4, respectively.) Note that at large distances it becomes increasingly difficult to have $G_1 \ll G_0$ because this requires extremely precise angular directivity in the antennas.

9. Under what conditions is the single-slope path loss model (2.24) the same as the free-space path loss model (2.7)?

10. Consider a receiver with noise power –160 dBm within the signal bandwidth of interest. Assume a single-slope path loss model with $d_0 = 1\ \text{m}, K$ obtained from the free-space path loss formula with isotropic antennas and $f_c = 1$ GHz, and $\gamma = 4$. For a transmit power of $P_t = 10\ \text{mW}$, find the maximum distance between the transmitter and receiver such that the received signal-to-noise power ratio is 20 dB.

11. For the set of empirical measurements of P_r/P_t given in Table 2.2, find the path loss exponent γ and parameter K that minimizes the MSE between the single-slope model (2.25) in dB and the empirical dB power measurements, assuming that $d_r = 1\,\mathrm{m}$. This problem differs from Example 2.5.1 in that now the two parameters γ and K are jointly optimized to minimize MSE, rather than fixing the value of K based on free-space path loss or the measured attenuation at the reference distance d_r. Find the received power at 150 m for this single-slope path loss model with transmit power of 1 mW (0 dBm) and compare with the result in Example 2.5.1. Does the better fitting of the data through this two-dimensional optimization lead to a larger or smaller value for this received power? Will this always be the case for any set of empirical measurements (why or why not)?

12. Find parameters for a multi-slope model with three segments to approximate the two-ray model path loss (2.14) over distances between 10 and 1000 meters, assuming $h_t = 10\,\mathrm{m}, h_r = 2\,\mathrm{m}$, and $G_0 = G_1 = 1$. Plot the path loss and the piecewise linear approximation using these parameters over this distance range.

13. This problem shows how different propagation models can lead to very different SNRs (and therefore different link performance) for a given system design. Consider a linear cellular system using frequency division, as might operate along a highway or rural road (see Figure 2.15). Each cell is allocated a certain band of frequencies, and these frequencies are reused in cells spaced a distance d away. Assume the system has square cells, 2 km per side, and that all mobiles transmit at the same power P. For the following propagation models, determine the minimum distance that the cells operating in the same frequency band must be spaced so that uplink SNR (the ratio of the minimum received signal-to-interference or S/I power from mobiles to the base station) is greater than 20 dB. You can ignore all interferers except those from the two nearest cells operating at the same frequency.

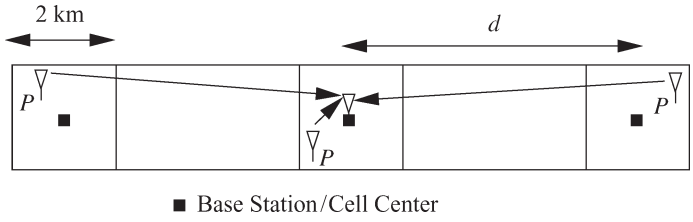


Figure 2.15: Linear cellular system for Problem 2-14.

   (a) Propagation for both signal and interference follow a free-space model.
   (b) Propagation for both signal and interference follow the single-slope path loss model (2.24) with $d_0 = 100\,\mathrm{m}, K = 1$, and $\gamma = 3$.
   (c) Propagation for the signal follows the single-slope path loss model with $d_0 = 100\,\mathrm{m}$, $K = 1$, and $\gamma = 2$, while propagation of the interference follows the same model but with $\gamma = 4$.

14. Show that the log-normal distribution for ψ given in (2.34) yields the Gaussian distribution for ψ_{dB} given in (2.37).

15. Table 2.7 lists a set of empirical path loss measurements. Assume a carrier frequency $f_c = 2$ GHz.

   (a) Find the parameters of a single-slope path loss model plus log-normal shadowing that best fit this data assuming K is calculated from free-space path loss at the reference distance $d_r = 1$ m.
   (b) Find the path loss at 2 km based on this model.

Table 2.7: path loss measurements for Problem 2-14

.

Distance from transmitter	P_r/P_t
5 m	–60 dB
25 m	–80 dB
65 m	–105 dB
110 m	–115 dB
400 m	–135 dB
1000 m	–150 dB

(c) Find the outage probability at a distance d assuming the received power at d due to path loss alone is 10 dB above the required power for non-outage.

16. Consider a cellular system operating at 900 MHz where propagation follows free-space path loss with variations about this path loss due to log-normal shadowing with standard deviation $\sigma = 6\,\mathrm{dB}$. Suppose that for acceptable voice quality a signal-to-noise power ratio of 15 dB is required at the mobile. Assume the base station transmits at 1 W and that its antenna has a 3 dB gain. There is no antenna gain at the mobile, and the receiver noise in the bandwidth of interest is $-60\ \mathrm{dBm}$. Find the maximum cell size such that a mobile on the cell boundary will have acceptable voice quality 90% of the time.

17. In this problem we will simulate the log-normal fading process over distance based on the autocovariance model (2.44). As described in the text, the simulation first generates a white noise process and then passes it through a first-order filter with a pole at $e^{-\delta/X_c}$. Assume $X_c = 20\,\mathrm{m}$ and plot the resulting log-normal fading process over a distance d ranging from $0\ \mathrm{m}$ to $200\ \mathrm{m}$, sampling the process every meter. You should normalize your plot about $0\ \mathrm{dB}$, since the mean of the log-normal shadowing is captured by path loss.

18. In this problem we will explore the impact of different log-normal shadowing parameters on outage probability. Consider a cellular system where the received signal power in dB has a Gaussian distribution with mean $P_r\ \mathrm{dBm}$ and standard deviation $\sigma\ \mathrm{dB}$. Assume the received signal power must be above 10 dBm for acceptable performance.

    (a) What is the outage probability when $P_r = 15\ \mathrm{dBm}$ and $\sigma = 8\ \mathrm{dB}$?

    (b) For $\sigma = 4\ \mathrm{dB}$, find the value of P_r required for the outage probability to be less than l% – a typical value for cellular systems.

    (c) Repeat part (b) for $\sigma = 12\ \mathrm{dB}$.

    (d) One proposed technique for reducing outage probability is to use *macrodiversity*, where a mobile unit's signal is received by multiple base stations and then combined. This can only be done if multiple base stations are able to receive a given mobile's signal. Explain why this might reduce outage probability.

19. Derive the formula for cell coverage percentage (2.55) by applying integration by parts to (2.53).

20. The cell coverage percentage is independent of the decorrelation distance X_c of the shadowing. Explain why. Now suppose it is known that a particular set of cell locations near the transmitter are in outage. Will the cell coverage percentage with this side information now depend on the decorrelation distance and, if so, how and why?

21. Find the coverage area for a microcellular system where path loss follows the single-slope model (with $\gamma = 3,\ d_0 = 1$ and $K = 0$ dB) and there is also log-normal shadowing with $\sigma = 4$ dB. Assume a cell radius of 100 m, a transmit power of 80 mW, and a minimum received power requirement of $P_{\min} = -100$ dBm.

22. Consider a cellular system where (a) path loss follows the single-slope model with $\gamma = 6$ and (b) there is also log-normal shadowing with $\sigma = 8$ dB. If the received power at the cell boundary due to path loss is 20 dB higher than the minimum required received power for non-outage, find the cell coverage area.

23. In microcells, path loss exponents usually range from 2 to 6 and shadowing standard deviation typically ranges from 4 to 12. Given a cellular system in which the received power due to path loss at the cell boundary equals the desired level for non-outage, find the path loss and shadowing parameters within these ranges that yield the best and worst coverage area. What is the coverage area when these parameters are in the middle of their typical ranges?

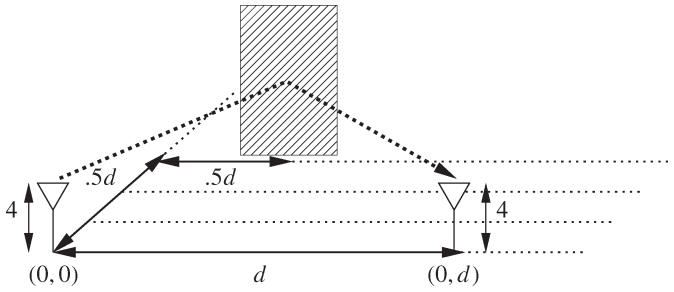


Figure 2.16: System with scattering for Problem 2-11.

24. What average power falloff with distance do you expect for the ten-ray model if $G_0 = G_1 >> G_i$ for $i = 2, ..., 9$? Repeat for $G_0 = G_1 = G_2 >> G_i$ for $i = 3, ..., 9$. Explain your answers in both cases.

25. For the ten-ray model, assume that the transmitter and receiver are at the same height in the middle of a street of width 20 m. The transmitter-receiver separation is 500 m. Find the delay spread for this model.

26. Consider a system with a transmitter, receiver, and scatterer as shown in Figure 2.16. Assume the transmitter and receiver are both at heights $h_t = h_r = 4$ m and are separated by distance d, with the scatterer at distance $.5d$ along both dimensions in a two-dimensional grid of the ground – that is, on such a grid the transmitter is located at $(0, 0)$, the receiver at $(0, d)$, and the scatterer at $(.5d,\ .5d)$. Assume a radar cross-section of 20 dBm^2, $G_s = 1$, and $f_c = 900$ MHz. Find the path loss of the scattered signal for $d = 1, 10, 100$, and 1000 meters. Compare with the path loss at these distances if the signal is only reflected, with reflection coefficient $R = -1$.

27. Find the median path loss under the Hata model assuming $f_c = 900$ MHz, $h_t = 20$ m, $h_r = 5$ m, and $d = 100$ m for a large urban city, a small urban city, a suburb, and a rural area. Explain qualitatively the path loss differences for these four environments.

28. Consider a wireless link operating outdoors over 200m in the mmW frequency band. Assume the path loss follows the single-slope model with a path loss exponent $\gamma = 2$ and reference distance for antenna far field $d_r = 1$ m. Without any additional attenuation caused by atmospheric and rain conditions, the model attenuation factor $K = 1$. Assume rain attenuation follows the simple approximation $K_{rain} = .95R^{.77}$ for R the rain rate in mm/hr. Assume the atmospheric absorption given in Fig. 2.13.

(a) What is the received signal power due to path loss and oxygen absorption at carrier frequencies of 60 GHz and of 80 GHz assuming a transmit power of 1W?

(b) Consider now only the 80 GHz link. Assume a day where it is dry at 8 am, there is a heavy drizzle of 2.5mm/Hr at 12 pm and a heavy downpour of 50mm/Hr at 5 pm. What is the required transmit power at 8 am, 12 pm and 5 pm if we desire a received signal power of -50dBm at each of these time instances?

29. Using the indoor attenuation model, determine the required transmit power for a desired received power of –110 dBm for a signal transmitted over 100 m that goes through three floors with attenuation 15 dB, 10 dB, and 6 dB (respectively) as well as two double plaster-board walls. Assume a reference distance $d_r = 1$, exponent $\gamma = 4$, and constant $K = 0$ dB.

30. Narrowband approximation in two-ray model: In this problem we examine the validity of the narrowband approximation in the two ray model. Consider the received signal for the two ray model (Equation (2.12)), where $u(t) = e^{j2\pi f_s t}$, $G_{r_0} = G_{r_1} = 1$, $G_{t_0} = G_{t_1} = 1, d = 10$ km, $h_t = h_r = 50$ m.

(a) Compute the delay spread τ. Assuming distance d is uncertain within $\pm 1\%$, what range does the delay spread lie in? (need not be exact, you can see this from a plot or by evaluating at some points)

(b) Compute the phase difference $\Delta\phi$ between the signal components in the following cases:

i. $f_s = \frac{1}{m}$ where m is the mid point of the range identified above
ii. $f_s = \frac{1}{2m}$
iii. $f_s = \frac{1}{100m}$

In which of the above cases does the approximation $u(t) \approx u(t - \tau)$ hold best (why)?

Hint: Phase "wraps" around at 2π and its integral multiples so the phase angles found in this problem should be reduced to lie in the range $(-\pi, \pi]$ — e.g. a phase difference of 4.5π should be reported as 0.5π.

Bibliography

[1] W. C. Jakes, Jr., *Microwave Mobile Communications*, Wiley, New York, 1974 [reprinted by IEEE Press].

[2] D. Parsons, *The Mobile Radio Propagation Channel*, 2nd Ed., Wiley, New York, 2000.

[3] M. Pätzold, *Mobile Fading Channels*, 2nd Ed., Wiley, New York, 2011.

[4] H. L. Bertoni, *Radio Propagation for Modern Wireless Systems*, Prentice Hall, 1999.

[5] J. Cavers, *Mobile Channel Characteristics*, Springer 2000.

[6] *IEEE J. Sel. Areas Commun.*, Special Issue on Channel and Propagation Modeling for Wireless Systems Design, April 2002 and August 2002.

[7] N. Costa and S. Haykin, *Multiple-input Multiple-output Channel Models: Theory and Practice*, John Wiley and Sons, 2010.

[8] *IEEE J. Sel. Areas Commun.*, Special Issue on Ultra-Wideband Radio in Multiaccess Wireless Communications, December 2002.

[9] T. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock. *Millimeter Wave Wireless Communications*. Pearson Education, 2014.

[10] G. H. Spencer and M. V. R. K. Murty, "General ray-tracing procedure", *Journal of the Optical Society of America*, Vol. 52, Issue 6, pp. 672-678, June 1962.

[11] T. S. Rappaport, *Wireless Communications-Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

[12] J. L. Volakis, *Antenna Engineering Handbook*, 4th Ed. McGraw-Hill, 2007.

[13] H. T. Friis, "A note on a simple transmission formula", *Proc. IRE*, vol. 34, no. 5, pp. 254-256, May 1946.

[14] A. S. Y. Poon and R. W. Brodersen, "The role of multiple-antenna systems in emerging open access environments," *EE Times Commun. Design Conf.*, October 2003.

[15] A. J. Rustako, Jr., N. Amitay, G. J. Owens, and R. S. Roman, "Radio propagation at microwave frequencies for line-of-sight microcellular mobile and personal communications," *IEEE Trans. Veh. Tech.*, pp. 203–10, February 1991.

[16] W. C. Y. Lee, *Mobile Communications Engineering*, McGraw-Hill, New York, 1982.

[17] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Sel. Areas Commun*., pp. 1205–11, July 1999.

[18] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Ant. Prop*., pp. 207–17, February 1992.

[19] A. F. Toledo and A. M. D. Turkmani, "Propagation into and within buildings at 900, 1800, and 2300 MHz," *Proc. IEEE Veh. Tech. Conf*., pp. 633–6, May 1992.

[20] F. C. Owen and C. D. Pudney, "Radio propagation for digital cordless telephones at 1700 MHz and 900 MHz," *Elec. Lett*., pp. 52–3, September 1988.

[21] A. F. Toledo, A. M. D. Turkmani, and J. D. Parsons, "Estimating coverage of radio transmission into and within buildings at 900, 1800, and 2300 MHz," *IEEE Pers. Commun. Mag*., pp. 40–7, Apri1 1998.

[22] D. M. J. Devasirvathan, R. R. Murray, and D. R. Woiter, "Time delay spread measurements in a wireless local loop test bed," *Proc. IEEE Veh. Tech. Conf*., pp. 241–5, May 1995.

[23] G. Durgin, T. S. Rappaport, and H. Xu, "Partition-based path loss analysis for in-home and residential areas at 5.85 GHz," *Proc. IEEE Globecom Conf*., pp. 904–9, November 1998.

[24] T. Kurner, D. J. Cichon, and W. Wiesbeck, "Concepts and results for 3D digital terrain-based wave propagation models: An overview," *IEEE J. Sel. Areas Commun*., pp. 1002–12, September 1993.

[25] A. J. Rustako, Jr., N. Amitay, G. J. Owens, and R.S. Roman, "Radio propagation at microwave frequencies for line-of-sight microcellular mobile and personal communications," *IEEE Trans. Veh. Technol*., vol. 40, pp. 203–210, February 1991.

[26] N. Amitay, "Modeling and computer simulation of wave propagation in lineal line-of-sight microcells," *IEEE Trans. Veh. Tech*., pp. 337–42, November 1992.

[27] J. W. McKown and R. L. Hamilton, Jr., "Ray tracing as a design tool for radio networks," *IEEE Network*, pp. 27–30, November 1991.

[28] K. A. Remley, H. R. Anderson, and A. Weisshar, "Improving the accuracy of ray-tracing techniques for indoor propagation modeling," *IEEE Trans. Veh. Tech*., pp. 2350–8, November 2000.

[29] A. Domazetovic, L. J. Greenstein, N. Mandayan, and I. Seskar, "A new modeling approach for wireless channels with predictable path geometries," *Proc. IEEE Veh. Tech. Conf*., September 2002.

[30] J. H. Tarng, W.-S. Liu, Y.-F. Huang, and J.-M. Huang, "A novel and efficient hybrid model of radio multipath-fading channels in indoor environments," *IEEE Trans. Ant. Prop*., pp. 585–94, March 2003.

[31] Z. Yun and M. F. Iskander. "Ray tracing for radio propagation modeling: Principles and applications." *IEEE Access*, Vol. 3, 2015, pp. 1089-1100.

[32] J.-F. Wagen, "Signal strength measurements at 881 MHz for urban microcells in downtown Tampa," *Proc. IEEE Globecom Conf*., pp. 1313–17, December 1991.

[33] R. J. C. Bultitude and G. K. Bedal, "Propagation characteristics on microcellular urban mobile radio channels at 910 MHz," *IEEE J. Sel. Areas Commun*., pp. 31–9, January 1989.

[34] J.-E. Berg, R. Bownds, and F. Lotse, "Path loss and fading models for microcells at 900 MHz," *Proc. IEEE Veh. Tech. Conf.*, pp. 666–71, May 1992.

[35] J. H. Whitteker, "Measurements of path loss at 910 MHz for proposed microcell urban mobile systems," *IEEE Trans. Veh. Tech.*, pp. 125–9, August 1988.

[36] H. Börjeson, C. Bergljung, and L. G. Olsson, "Outdoor microcell measurements at 1700 MHz," *Proc. IEEE Veh. Tech. Conf.*, pp. 927–31, May 1992.

[37] K. Schaubach, N. J. Davis IV, and T. S. Rappaport, "A ray tracing method for predicting path loss and delay spread in microcellular environments," *Proc. IEEE Veh. Tech. Conf.*, pp. 932–5, May 1992.

[38] F. Ikegami, S. Takeuchi, and S. Yoshida, "Theoretical prediction of mean field strength for urban mobile radio," *IEEE Trans. Ant. Prop.*, pp. 299–302, March 1991.

[39] M. C. Lawton and J. P. McGeehan, "The application of GTD and ray launching techniques to channel modeling for cordless radio systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 125–30, May 1992.

[40] J. B. Keller, "Geometrical theory of diffraction," *J. Opt. Soc. Amer.*, 52, pp. 116–30, 1962.

[41] R. J. Luebbers, "Finite conductivity uniform GTD versus knife edge diffraction in prediction of propagation path loss," *IEEE Trans. Ant. Prop.*, pp. 70–6, January 1984.

[42] C. Bergljung and L. G. Olsson, "Rigorous diffraction theory applied to street microcell propagation," *Proc. IEEE Globecom Conf.*, pp. 1292–6, December 1991.

[43] R. G. Kouyoumjian and P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface," *Proc. IEEE*, pp. 1448–61, November 1974.

[44] G. K. Chan, "Propagation and coverage prediction for cellular radio systems," *IEEE Trans. Veh. Tech.*, pp. 665–70, November 1991.

[45] K. C. Chamberlin and R. J. Luebbers, "An evaluation of Longley–Rice and GTD propagation models," *IEEE Trans. Ant. Prop.*, pp. 1093–8, November 1982.

[46] J. Walfisch and H. L. Bertoni, "A theoretical model of UHF propagation in urban environments," *IEEE Trans. Ant. Prop.*, pp. 1788–96, October 1988.

[47] D. Chizhik, G. J. Foschini, M. J. Gans and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 2, pp. 361-368, April 2002.

[48] W. Zhang, "Fast two-dimensional diffraction modeling for site-specific propagation prediction in urban microcellular environments," *IEEE Trans. Veh. Techn.*, vol. 49, no. 2, pp. 428-436, 2000.

[49] S. Deng, G. R. MacCartney and T. S. Rappaport, "Indoor and outdoor 5G diffraction measurements and models at 10, 20, and 26 GHz," *Proc. 2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-7.

[50] M. Jacob, S. Priebe, R. Dickhoff, T. Kleine-Ostmann, T. Schrader and T. Kurner, "Diffraction in mm and Sub-mm Wave indoor propagation channels," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 833-844, March 2012.

[51] M. I. Skolnik, *Introduction to Radar Systems*, 2nd ed., McGraw-Hill, New York, 1980.

[52] S. Y. Seidel, T. S. Rappaport, S. Jain, M. L. Lord, and R. Singh, "Path loss, scattering, and multipath delay statistics in four European cities for digital cellular and microcellular radiotelephone," *IEEE Trans. Veh. Tech.*, pp. 721–30, November 1991.

[53] S. T. S. Chia, "1700 MHz urban microcells and their coverage into buildings," *Proc. IEEE Ant. Prop. Conf.*, pp. 504–11, York, U.K., Apri1 1991.

[54] T. Fugen, J. Maurer, W. Sorgel and W. Wiesbeck, "Characterization of multipath clusters with ray-tracing in urban MIMO propagation environments at 2 GHz," *in 2005 IEEE Antennas and Propagation Society International Symposium*, Washington, DC, 2005, pp. 410-413.

[55] M. A. Jensen and J. W. Wallace, "A review of antennas and propagation for MIMO wireless communications," *IEEE Transactions on Antennas and Propagation*, pp. 2810-2824, Nov. 2004.

[56] C. Pan et al., "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14-20, June 2021.

[57] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding and M. Gidlund, "Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6G wireless networks," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 184-191, December 2021.

[58] K. H. Ng, E. K. Tameh, A. Doufexi, M. Hunukumbure and A. R. Nix, "Efficient multielement ray tracing with site-specific comparisons using measured MIMO channel data," *IEEE Transactions on Vehicular Technology*, pp. 1019-1032, May 2007.

[59] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 38.901, March 2017, Website: www.3gpp.org/DynaReport/38901.htm

[60] IMT, "Guidelines for evaluation of radio interface technologies for IMT-2020," Tech Rep. ITU-R M.2412-0, Oct. 2017, Website: www.itu.int/pub/R-REP-M.2412.

[61] P. Kyosti et al., "WINNER II channel models," Eur. Commission, IST-4-027756-WINNER, Brussels, Belgium, Tech. Rep. D1.1.2, Sep. 2007

[62] A. A. M. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128-137, February 1987.

[63] L. Schumacher, K. I. Pedersen, and P.E. Mogensen, "From antenna spacings to theoretical capacities - guidelines for simulating MIMO systems," *Proc. Personal, Indoor and Mobile Radio Communications (PIMRC) Conf.*, Sept. 2002, pp. 587-592.

[64] V. Erceg et al., "TGn Channel Models." Doc. IEEE802.11-03/940r4.

[65] G. Breit et. al., "TGac Channel Model Addendum." IEEE 802.11-09/0308r12.

[66] J. Liu et. al., "IEEE 802.11ax Channel Model Document." IEEE 802.11-14/0882r4.

[67] FCC Memorandum and Opinion Order 05-45 *WTB Allocations and Service Rules for the 71-76 GHz, 81-86 GHz and 92-95 GHz Bands.* Feb. 2005.

[68] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, pp. 101-107, June 2011.

[69] V. Raghavan, J. Li, A. Sampath, O. H. Koymen, and T. Luo, *Millimeter Wave Communications in 5G and Towards 6G*, CRC Press, Boca Raton, FL., 2025.

[70] M. Marcus and B. Pattan, "Millimeter wave propagation: Spectrum management implications," *IEEE Microwave Mag.*, pp. 54-62, June 2005.

[71] ITU-R P.838-3, "Specific Model for Rain for Use in Prediction Methods," 2005.

[72] J. Ostrometzky and H. Messer, "Accumulated rainfall estimation using maximum attenuation of microwave radio signal," *Proc. IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 193-196, June 2014.

[73] M. Xiao et al., "Millimeter Wave Communications for Future Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909-1935, Sept. 2017.

[74] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks - with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213-6230, Dec. 2017.

[75] M. Shafi, J. Zhang, H. Tataria, A. F. Molisch, S. Sun, T. S. Rappaport, F. Tufvesson, S. Wu, K. Kitao, "Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems," *IEEE Communications Magazine*, Vol. 56, No. 12., pp. 14-20, Dec. 2018.

[76] I. F. Akyildiza, J. M. Jornet, C. Han, "Terahertz band: Next frontier for wireless communications," *Elsevier J. Physical Commun.*, Volume 12, pp. 16-21, Sept. 2014.

[77] METIS Channel Model METIS2020, Deliverable D1.4 v3. [Online]. Available: https://www.metis2020.com/wpcontent/uploads/deliverables/METIS_D1.4_v1.0.pdf. July 2015.

[78] E. Ostlin, H.-J. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Techn.*, vol. 59, no. 6, pp. 2735-2747, 2010.

[79] P. Almers et. al., "Survey of channel and radio propagation models for wireless MIMO Systems," *EURASIP J. Wireless Commun. Net.*, vol. 2007.

[80] J. Medbo et al., "Channel modeling for the fifth generation mobile communications," Proc. 8th European Conference on Antennas and Propagation (EuCAP 2014), The Hague, 2014, pp. 219-223.

[81] P. Ferrand, M. Amara, S. Valentin and M. Guillaud, "Trends and challenges in wireless channel modeling for evolving radio access," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 93-99, July 2016.

[82] S. Sun, T. Rappaport, M. Shafi, P. Tang, J. Zhang, and P.J. Smith, "Propagation Models and Performance Evaluation for 5G Millimeter-Wave Bands," *IEEE Transactions on Vehicular Technology*, Vol. 67 , No. 9 , Sept. 2018.

[83] M. Kasparick, R. L. G. Cavalcante, S. Valentin, S. Stanczak, and M. Yukawa, "Kernel-based adaptive online reconstruction of coverage maps with side information," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5461–5473, 2016.

[84] D. Romero, S.-J. Kim, G. B. Giannakis, and R. Lopez-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process*., vol. 65, no. 10, pp. 2547–2560, 2017.

[85] R. Nikbakht, A. Jonsson, and A. Lozano, "Dual-kernel online reconstruction of power maps," *IEEE Global Commun. Conf. (GLOBECOM18)*, 2018.

[86] D. Wong and D. C. Cox, "Estimating local mean signal power level in a Rayleigh fading environment," *IEEE Trans. Veh. Tech*., pp. 956–9, May 1999.

[87] T. Okumura, E. Ohmori, and K. Fukuda, "Field strength and its variability in VHF and UHF land mobile service," *Rev. Elec. Commun. Lab*., pp. 825–73, September/October 1968.

[88] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Tech*., pp. 317-325, August 1980.

[89] European Cooperative in the Field of Science and Technical Research EURO-COST 231, "Urban transmission loss models for mobile radio in the 900 and 1800 MHz bands," rev. 2, The Hague, September 1991.

[90] A. F. Molisch, *Wireless Communications: From Fundamentals to Beyond 5G*, Wiley-IEEE Press, 3rd ed. 2022.

[91] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The COST 259 directional channel model-part I: Overview and methodology," *IEEE Trans. on Wireless Communications*, vol. 5, no. 12, pp. 3421 -3433, Dec. 2006.

[92] N. Czink and C. Oestges, "The COST 273 MIMO Channel Model: Three Kinds of Clusters," *IEEE 10th International Symposium on Spread Spectrum Techniques and Applications*, pp. 282-286, Aug. 2008.

[93] L. Liu et al., "The COST 2100 MIMO channel model," *IEEE Wireless Communications*, pp. 92-99, Dec. 2012.

[94] E. McCune and K. Feher, "Closed-form propagation model combining one or more propagation constant segments," *Proc. IEEE Veh. Tech. Conf*., pp. 1108–12, May 1997.

[95] M. Feuerstein, K. Blackard, T. Rappaport, S. Seidel, and H. Xia, "Path loss, delay spread, and outage models as functions of antenna height for microcellular system design," *IEEE Trans. Veh. Tech*., pp. 487–98, August 1994.

[96] P. Harley, "Short distance attenuation measurements at 900 MHz and 1.8 GHz using low antenna heights for microcells," *IEEE J. Sel. Areas Commun*., pp. 5–11, January 1989.

[97] A. J. Motley and J. M. P. Keenan, "Personal communication radio coverage in buildings at 900 MHz and 1700 MHz," *Elec. Lett*., pp. 763–4, June 1988.

[98] S. Y. Seidel, T. S. Rappaport, M. J. Feuerstein, K. L. Blackard, and L. Grindstaff, "The impact of surrounding buildings on propagation for wireless in-building personal communications system design," *Proc. IEEE Veh. Tech. Conf*., pp. 814–18, May 1992.

[99] C. R. Anderson, T. S. Rappaport, K. Bae, A. Verstak, N. Tamakrishnan, W. Tranter, C. Shaffer, and L. T. Waton, "In-building wideband multipath characteristics at 2.5 and 60 GHz," *Proc. IEEE Veh. Tech. Conf*., pp. 24–8, September 2002.

[100] L.-S. Poon and H.-S. Wang, "Propagation characteristic measurement and frequency reuse planning in an office building," *Proc. IEEE Veh. Tech. Conf.*, pp. 1807–10, June 1994.

[101] A. Ghosh et al., "Millimeter-wave enhanced local area systems: A high data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, pp. 1152–1163, Jun. 2014.

[102] X. Zhang and J.G. Andrews, "Downlink cellular network analysis with multi-slope path loss models," *IEEE Trans. Commun.*, pp. 1881-1894, May 2015.

[103] R. Hoppe, G. Wölfle, and F. M. Landstorfer, "Measurement of building penetration loss and propagation models for radio transmission into buildings," *Proc. IEEE Veh. Tech. Conf.*, pp. 2298–2302, Apri1 1999.

[104] E. H. Walker, "Penetration of radio signals into buildings in cellular radio environments," *Bell Systems Tech. J.*, pp. 2719–34, September 1983.

[105] S. S. Ghassemzadeh, L. J. Greenstein, A. Kavcic, T. Sveinsson, and V. Tarokh, "Indoor path loss model for residential and commercial buildings," *Proc. IEEE Veh. Tech. Conf.*, pp. 3115–19, October 2003.

[106] A. J. Goldsmith, L. J. Greenstein, and G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing," *IEEE Trans. Veh. Tech.*, pp. 439–46, August 1994.

[107] A. J. Goldsmith and L. J. Greenstein, "A measurement-based model for predicting coverage areas of urban microcells," *IEEE J. Sel. Areas Commun.*, pp. 1013–23, September 1993.

[108] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Elec. Lett.*, pp. 2145–6, November 7, 1991.

[109] G. L. Stüber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[110] A. Algans, K. I. Pedersen, and P. E. Mogensen, "Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading," *IEEE J. Sel. Areas Commun.*, pp. 523–31, April 2002.

[111] M. Marsan and G. C. Hess, "Shadow variability in an urban land mobile radio environment," *Elec. Lett.*, pp. 646–8, May 1990.

[112] J. Weitzen and T. Lowe, "Measurement of angular and distance correlation properties of lognormal shadowing at 1900 MHz and its application to design of PCS systems," *IEEE Trans.Veh. Tech.*, pp. 265–73, March 2002.

[113] W. Turin, R. Jana, S. S. Ghassemzadeh, V.W. Rice, and V. Tarokh, "Autoregressive modeling of an indoor UWB channel," *Proc. IEEE Conf. UWB Syst. Technol.*, pp. 71–4, May 2002.

Chapter 3

Statistical Multipath Channel Models

In this chapter we examine fading models for the constructive and destructive addition of different multipath components introduced by the channel. Although these multipath effects are captured in the ray-tracing models from Chapter 2 for deterministic channels, we rarely have sufficiently precise knowledge about wireless propagation environments to characterize them deterministically. Thus wireless channels are often characterized statistically. The statistical characterization of the random multipath channel is based on its time-varying impulse response. This chapter develops a statistical characterization of this channel model and describes its important properties.

If a single pulse is transmitted over a multipath channel then the received signal is a pulse train, with each pulse in the train corresponding to the LOS component or a multipath component associated with a distinct reflector. The *delay spread* of a multipath channel characterizes the difference in arrival times of the first received signal component (LOS or, if blocked, the first received multipath component) and the last received signal component of significant energy associated with a single transmitted pulse. For multi-ray reflection models such as the 10-ray and 6-ray models of Chapter 2.8.1, the delay spread equals the delay difference between the first and last ray. The delay spread for random channels is characterized statistically, as described in 3.3.2. If the delay spread is small compared to the inverse of the signal bandwidth, then there is little time spreading in the received signal. However, if the delay spread is large then there is significant time spreading of the received signal, which leads to substantial signal distortion.

Another characteristic of the multipath channel is its time-varying nature. This time variation arises because either the transmitter, receiver, or objects in the environment are moving. Hence the location of reflectors in the transmission path, which gives rise to multipath, will change over time. Thus, if we repeatedly transmit pulses from a transmitter given these changing reflector locations, we will observe changes in the amplitudes, delays, and number of multipath components corresponding to each pulse. However, these changes occur over a much larger time scale than the fading due to constructive and destructive addition of multipath components associated with a fixed set of scatterer locations. We will first describe the time-varying channel impulse response that models these changing multipath components. We then restrict this model to narrowband fading, where the channel bandwidth is small compared to the inverse delay spread. For this narrowband model we will assume a quasi-static environment featuring a fixed number of multipath components, each with fixed path loss and shadowing. For this environment we then characterize the variations over short distances (small-scale variations) due to the constructive and destructive addition of multipath components. We also characterize the statistics of wideband multipath channels using two-dimensional transforms based on the underlying time-varying impulse response. Discrete-time and MIMO channel models are also discussed.

3.1 Time-Varying Channel Impulse Response

Let the transmitted signal be as in (2.1):

$$s(t) = \mathrm{Re}\{u(t)e^{j2\pi f_c t}\} = \mathrm{Re}\{u(t)\}\cos(2\pi f_c t) - \mathrm{Im}\{u(t)\}\sin(2\pi f_c t), \tag{3.1}$$

where $u(t)$ is the equivalent lowpass signal for $s(t)$ with bandwidth B_u and f_c is its carrier frequency. Neglecting noise, the corresponding received signal is the sum of the LOS path and all *resolvable* multipath components:

$$r(t) = \mathrm{Re}\left\{\sum_{i=0}^{N(t)-1} \alpha_i(t)u(t-\tau_i(t))e^{j(2\pi f_c(t-\tau_i(t))+\phi_{D_i}(t))}\right\}, \tag{3.2}$$

where the first term in the summation, corresponding to $i = 0$, is associated with the LOS path. The unknowns in this expression are the number of *resolvable* signal components $N(t)$ and, for each signal component (LOS and multipath), its path length $x_i(t)$ and corresponding delay $\tau_i(t) = x_i(t)/c$, Doppler phase shift $\phi_{D_i}(t)$, and amplitude $\alpha_i(t)$. The phase change associated with user movement and the resolvability of different signal components will be discussed in the next two paragraphs. If at time t the LOS path is blocked or significantly attenuated then this will be captured in the amplitude term $\alpha_0(t)$. Similarly, the amplitude $\alpha_i(t)$ of the ith multipath component is determined by the path loss and shadowing along its propagation path. In addition, the amplitudes $\alpha_i(t), i = 0, \ldots, N(t) - 1$ may experience multipath fading due to constructive and destructive addition of nonresolvable multipath components.

When the transmitter or receiver is moving, the delay $\tau_i(t)$ in (3.2) changes over time and hence so does the associated phase change. This change is captured in (3.2) as $e^{-j2\pi f_c \tau_i(t)}$, however $\tau_i(t)$ is rarely precisely known in practice. Alternatively the phase change due to small changes in $\tau_i(t)$ can be captured by the signal's Doppler shift $f_{D_i}(t) = v(t)\cos\theta_i(t)/\lambda$ for $\theta_i(t)$ the component's angle of arrival (AoA) relative to the direction of motion at velocity $v(t)$. This Doppler frequency shift, which is generally easy to measure in practice, leads to a Doppler phase shift of $\phi_{D_i}(t) = \int_0^t 2\pi f_{D_i}(\nu)d\nu$. Note that the phase change of the ith multipath component due to user movement in (3.2) should either be captured in $e^{-j2\pi f_c \tau_i(t)}$ or in $\phi_{D_i}(t)$, but not in both terms as then the phase shift due to user movement is double-counted. It is generally assumed that $\tau_i(t) = x_i(t)/c$ does not capture changes in $x_i(t)$ of less than the signal wavelength λ, so the phase changes due to these small user movements are captured by the Doppler term in (3.2).

We now define resolvability of multipath components. We say that two multipath components with delay τ_i and τ_j are *resolvable* if their delay difference significantly exceeds the inverse signal bandwidth: $|\tau_i - \tau_j| \gg B_u^{-1}$. Multipath components that do not satisfy this resolvability criteria cannot be separated out at the receiver because $u(t - \tau_i) \approx u(t - \tau_j)$, and thus these components are *nonresolvable*. These nonresolvable multipath components are summed together to create a single resolvable multipath component in (3.2) corresponding to delay τ_i, with a resulting amplitude and phase corresponding to the summed signal. Any multipath components with delay τ_j for which $|\tau_j - \tau_0| < B_u^{-1}$ are combined with the LOS signal. The amplitudes of the LOS and resolvable multipath components in (3.2) that include nonresolvable multipath components will typically undergo rapid multipath fading due to the constructive and destructive combining of the different signal components. For wideband channels (B_u relatively large) each term in the summation of (3.2) typically corresponds to a single (smooth or non-smooth[1]) reflector, or to a cluster of closely-spaced reflectors, as shown in Figure 3.1. By contrast, narrowband channels (B_u relatively small) are likely to have one or more of the signal components in (3.2) correspond to a summation of nonresolvable signal components.

[1] A non-smooth reflector has a surface that is sufficiently rough such that it generates different multipath components with slightly different delays.

Since the parameters $\alpha_i(t), \tau_i(t)$, and $\phi_{D_i}(t)$ associated with each resolvable multipath component change over time, they are characterized as random processes that we assume to be both stationary and ergodic. Thus, the received signal is also a stationary and ergodic random process. For wideband channels, where each term in (3.2) corresponds to a single reflector, these parameters change slowly as the propagation environment changes. For narrowband channels, where each term in (3.2) results from the sum of nonresolvable multipath components, the parameters can change quickly – on the order of a signal wavelength – owing to constructive and destructive addition of the different components.

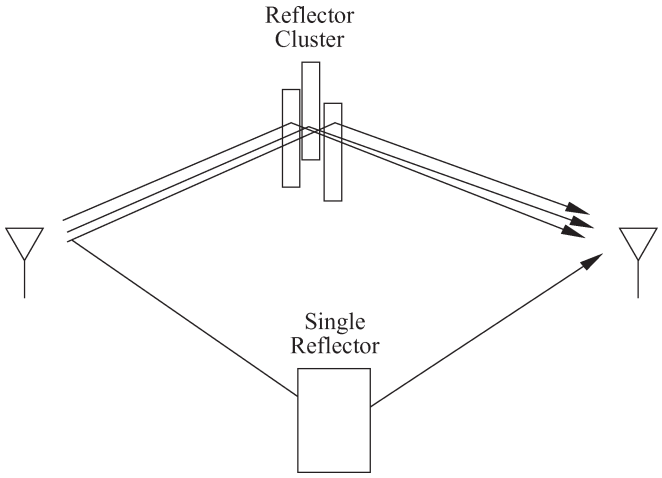


Figure 3.1: A single reflector and a cluster of closely-spaced reflectors. The reflector cluster generates multipath components that are typically resolvable in wideband channels and non-resolvable in narrowband channels.

We can simplify $r(t)$ by letting

$$\phi_i(t) = 2\pi f_c \tau_i(t) - \phi_{D_i}(t). \tag{3.3}$$

Then the received signal can be rewritten as

$$r(t) = \mathrm{Re}\left\{ \left[\sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} u(t - \tau_i(t)) \right] e^{j2\pi f_c t} \right\}. \tag{3.4}$$

Since $\alpha_i(t)$ is a function of path loss and shadowing while $\phi_i(t)$ depends on delay and Doppler, we typically assume that these two random processes are independent.

We denote the equivalent lowpass time-varying impulse response of the channel at time t to an impulse at time $t - \tau$ by $c(\tau, t)$. The received signal $r(t)$ is obtained by convolving the equivalent lowpass input signal $u(t)$ with $c(\tau, t)$ and then upconverting to the carrier frequency: [2]

$$r(t) = \mathrm{Re}\left\{ \left(\int_{-\infty}^{\infty} c(\tau, t) u(t - \tau) d\tau \right) e^{j2\pi f_c t} \right\}. \tag{3.5}$$

Note that $c(\tau, t)$ has two time parameters: the time t when the impulse response is observed at the receiver, and the time $t - \tau$ when the impulse is launched into the channel relative to the observation time t. If at time t there is no physical reflector in the channel with multipath delay $\tau_i(t) = \tau$, then $c(\tau, t) = 0$. Although the definition of the time-varying channel impulse response might at first seem counterintuitive, $c(\tau, t)$ must be defined in this way

[2] See Appendix A for discussion of the equivalent lowpass representation for bandpass signals and systems.

to be consistent with the special case of time-invariant channels. Specifically, for time-invariant channels we have $c(\tau, t) = c(\tau, t + T)$ for any T; that is, the response at time t to an impulse at time $t - \tau$ equals the response at time $t + T$ to an impulse at time $t + T - \tau$. Setting $T = -t$, we get that $c(\tau, t) = c(\tau, t - t) = c(\tau)$, where $c(\tau)$ is the standard time-invariant channel impulse response: the response at time τ to an impulse at time zero.[3]

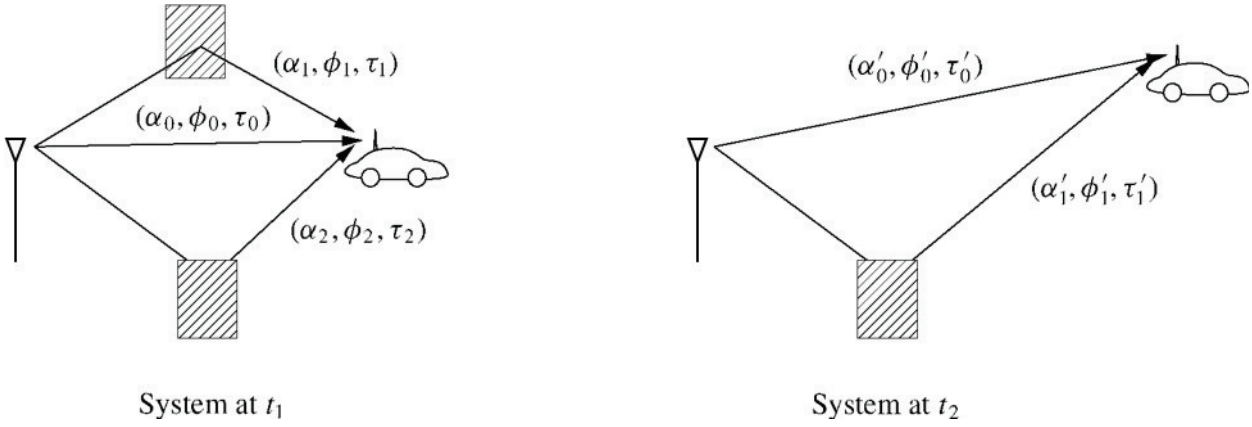


Figure 3.2: System multipath at two different measurement times.

We see from (3.4) and (3.5) that $c(\tau, t)$ must be given by

$$c(\tau, t) = \sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} \delta(\tau - \tau_i(t)). \tag{3.6}$$

Substituting (3.6) back into (3.5) yields (3.4), thereby confirming that (3.6) is the channel's equivalent lowpass time-varying impulse response:

$$\begin{aligned} r(t) &= \mathrm{Re}\left\{ \left[\int_{-\infty}^{\infty} c(\tau, t) u(t - \tau) d\tau \right] e^{j2\pi f_c t} \right\} \\ &= \mathrm{Re}\left\{ \left[\int_{-\infty}^{\infty} \sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} \delta(\tau - \tau_i(t)) u(t - \tau) d\tau \right] e^{j2\pi f_c t} \right\} \\ &= \mathrm{Re}\left\{ \left[\sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} \left(\int_{-\infty}^{\infty} \delta(\tau - \tau_i(t)) u(t - \tau) d\tau \right) \right] e^{j2\pi f_c t} \right\} \\ &= \mathrm{Re}\left\{ \left[\sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} u(t - \tau_i(t)) \right] e^{j2\pi f_c t} \right\}, \end{aligned} \tag{3.7}$$

where the last equality follows from the sifting property of delta functions:

$$\int_{-\infty}^{\infty} \delta(\tau - \tau_i(t)) u(t - \tau) d\tau = \delta(t - \tau_i(t)) * u(t) = u(t - \tau_i(t)). \tag{3.8}$$

[3]By definition, $c(\tau, 0)$ is the response at time zero to an impulse at time $-\tau$, but since the channel is time invariant, this equals the response at time τ to an impulse at time zero.

Some channel models assume a continuum of multipath delays, in which case the sum in (3.6) becomes an integral that simplifies to a time-varying complex amplitude associated with each multipath delay τ:

$$c(\tau, t) = \int_{-\infty}^{\infty} \alpha(\xi, t) e^{-j\phi(\xi,t)} \delta(\tau - \xi) d\xi = \alpha(\tau, t) e^{-j\phi(\tau,t)}. \tag{3.9}$$

For a concrete example of a time-varying impulse response, consider the system shown in Figure 3.2, where each multipath component corresponds to a single reflector. At time t_1 there are three multipath components associated with the received signal, each with amplitude, phase, and delay triple $(\alpha_i, \phi_i, \tau_i), i = 0, 1, 2$. Thus, impulses that were launched into the channel at time $t_1 - \tau_i, i = 0, 1, 2$, will all be received at time t_1, and impulses launched into the channel at any other time will not be received at t_1 (because there is no multipath component with the corresponding delay). The time-varying impulse response corresponding to t_1 equals

$$c(\tau, t_1) = \sum_{i=0}^{2} \alpha_i e^{-j\phi_i} \delta(\tau - \tau_i), \tag{3.10}$$

and the channel impulse response for $t = t_1$ is shown in Figure 3.3. Figure 3.2 also shows the system at time t_2, where there are two multipath components associated with the received signal having amplitude, phase, and delay triple $(\alpha_i', \phi_i', \tau_i'), i = 0, 1$. Thus, impulses that were launched into the channel at time $t_2 - \tau_i', i = 0, 1$, will all be received at time t_2, and impulses launched into the channel at any other time will not be received at t_2. The time-varying impulse response at t_2 equals

$$c(\tau, t_2) = \sum_{i=0}^{1} \alpha_i' e^{-j\phi_i'} \delta(\tau - \tau_i') \tag{3.11}$$

and is also shown in Figure 3.3.

If the channel is time invariant, then its response to an impulse at time t_1 is just a shifted version of its response to an impulse at time $t_2 \neq t_1$. Hence, the time-varying parameters in $c(\tau, t)$ become constant and $c(\tau, t) = c(\tau)$ is just a function of τ:

$$c(\tau) = \sum_{i=0}^{N} \alpha_i e^{-j\phi_i} \delta(\tau - \tau_i). \tag{3.12}$$

Example 3.1: Consider a wireless LAN operating in a factory near a conveyor belt. The transmitter and receiver have a LOS path between them with gain α_0, phase ϕ_0, and delay τ_0. Every T_0 seconds, a metal item comes down the conveyor belt, creating an additional reflected signal path with gain α_1, phase ϕ_1, and delay τ_1. Find the time-varying impulse response $c(\tau, t)$ of this channel.

Solution: For $t \neq nT_0, n = 1, 2, \ldots$, the channel impulse response simply corresponds to the LOS path. For $t = nT_0$, the channel impulse response includes both the LOS and reflected paths. Thus, $c(\tau, t)$ is given by

$$c(\tau, t) = \begin{cases} \alpha_0 e^{j\phi_0} \delta(\tau - \tau_0), & t \neq nT_0, \\ \alpha_0 e^{j\phi_0} \delta(\tau - \tau_0) + \alpha_1 e^{j\phi_1} \delta(\tau - \tau_1), & t = nT_0. \end{cases} \tag{3.13}$$

Note that, for typical carrier frequencies, the ith multipath component will have $f_c \tau_i(t) \gg 1$. For example, with $f_c = 1$ GHz and $\tau_i = 50$ ns (a typical value for an indoor system), $f_c \tau_i = 50 \gg 1$. Outdoor wireless systems

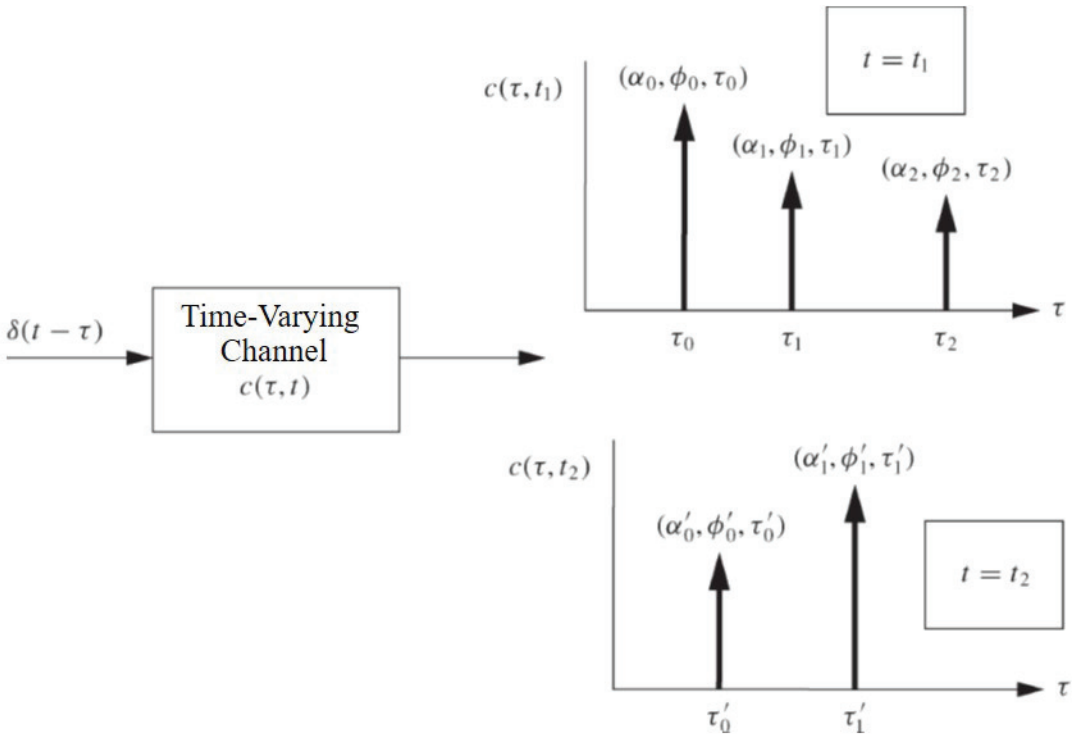


Figure 3.3: Response of time-varying channel.

have multipath delays much greater than 50 ns, so this property also holds for these systems. If $f_c\tau_i(t) \gg 1$ then a small change in the path delay $\tau_i(t)$ can lead to a large phase change in the ith multipath component with phase $\phi_i(t) = 2\pi f_c\tau_i(t) - \phi_{D_i}(t) - \phi_0$. Rapid phase changes in each multipath component give rise to constructive and destructive addition of the multipath components constituting the received signal, which in turn causes rapid variation in the received signal strength. This multipath fading phenomenon was exhibited in the deterministic channel models of Chapter 2 (see, e.g., Figure 2.4 illustrating the received signal power for the two-ray multipath model). This phenomenon for statistical multipath models will be discussed in more detail in subsequent sections.

The impact of multipath on the received signal depends on whether the spread of time delays associated with the LOS and different multipath components is large or small relative to the inverse signal bandwidth. If this channel delay spread is small then the LOS and all multipath components are nonresolvable, leading to the narrowband fading model described in the next section. If the delay spread is large then some number of the LOS and multipath components are resolvable, leading to the wideband fading model of Section 3.3. The delay spread is typically measured relative to the received signal component to which the demodulator is synchronized. Thus, for the time-invariant channel model of (3.12), if the demodulator synchronizes to the LOS signal component, which has the smallest delay τ_0, then the delay spread is a constant given by $T_m = \max_i[\tau_i - \tau_0]$. However, if the demodulator synchronizes to a multipath component with delay equal to the mean delay $\bar{\tau}$, then the delay spread is given by $T_m = \max_i |\tau_i - \bar{\tau}|$. In time-varying channels the multipath delays vary with time, so the delay spread T_m becomes a random variable. Moreover, some received multipath components have significantly lower power than others, so it's not clear how the delay associated with such components should be used in the characterization of delay spread. In particular, if the power of a multipath component is below the noise floor then it should not significantly contribute to the delay spread. These issues are typically dealt with by characterizing the delay spread relative to the channel power delay profile, defined in Section 3.3.2. Specifically, two common

characterizations of channel delay spread – average delay spread and rms (root mean square) delay spread – are determined from the power delay profile. Other characterizations of delay spread, such as excess delay spread, the delay window, and the delay interval, are sometimes used as well (see [1, Chap. 5.4.1] and [2, Chap. 6.7.1]). The exact characterization of delay spread is not that important for understanding the general impact of delay spread on multipath channels, as long as the characterization roughly measures the delay associated with significant multipath components. We will adopt the rms delay spread for our characterization as it is the one most widely used in the literature. Assuming the demodulator synchronizes to a signal component at the average delay spread, the rms delay spread is a good measure of the variation about this average. Channel delay spread is highly dependent on the propagation environment. In indoor channels delay spread typically ranges from 10 to 1000 nanoseconds, in urban microcells it ranges from 1–10 microseconds, and in rural macrocells it ranges from 10–100 microseconds [1, Chap. 5].

3.2 Narrowband Fading Model

Suppose the delay spread T_m of a channel is small relative to the inverse baseband signal bandwidth B_u of the transmitted signal; that is, suppose $T_m \ll B_u^{-1}$. We refer to the fading in this scenario as *narrowband fading* which is a function of both the signal bandwidth B_u and the channel delay spread T_m. As discussed previously, the delay spread T_m for time-varying channels is usually characterized by the rms delay spread, but it can also be characterized in other ways. Under most delay spread characterizations, $T_m \ll B_u^{-1}$ implies that the delay associated with the ith multipath component $\tau_i \leq T_m$ for all i, so $u(t-\tau_i) \approx u(t)$ for all i. Since in the narrowband fading model the delay associated with all multipath components results in minimal signal distortion in time, we say the channel under this model is *nondispersive*. We can rewrite (3.4) as

$$r(t) = \mathrm{Re}\left\{ u(t)e^{j2\pi f_c t} \left(\sum_i \alpha_i(t) e^{-j\phi_i(t)} \right) \right\}. \tag{3.14}$$

Equation (3.14) differs from the original transmitted signal by the complex scale factor in large parentheses. This scale factor is independent of the transmitted signal $s(t)$ and, in particular, of the equivalent lowpass signal $u(t)$ as long as the narrowband assumption $T_m \ll 1/B_u$ is satisfied. In order to characterize the random scale factor caused by the multipath, we choose $s(t)$ to be an unmodulated carrier with random phase offset ϕ_0:

$$s(t) = \mathrm{Re}\{e^{j(2\pi f_c t + \phi_0)}\} = \cos(2\pi f_c t + \phi_0), \tag{3.15}$$

which is narrowband for any T_m.

With this assumption the received signal becomes

$$r(t) = \mathrm{Re}\left\{ \left[\sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} \right] e^{j2\pi f_c t} \right\} = r_I(t)\cos(2\pi f_c t) - r_Q(t)\sin(2\pi f_c t), \tag{3.16}$$

where the in-phase and quadrature components are given by

$$r_I(t) = \sum_{i=0}^{N(t)-1} \alpha_i(t)\cos\phi_i(t), \tag{3.17}$$

$$r_Q(t) = \sum_{i=0}^{N(t)-1} \alpha_i(t)\sin\phi_i(t) \tag{3.18}$$

and where the phase term

$$\phi_i(t) = 2\pi f_c \tau_i(t) - \phi_{D_i}(t) - \phi_0 \tag{3.19}$$

now incorporates the phase offset ϕ_0 as well as the effects of delay and Doppler.

Suppose $N(t)$ is large and all signal (multipath as well as the LOS) components have random amplitudes with approximately the same distribution, and similarly for the random phases of each component. Then, given that $\alpha_i(t)$ and $\phi_i(t)$ are also independent for different components, we can invoke the Central Limit Theorem (CLT) for sums of independent identically-distributed (i.i.d.) random variables to approximate $r_I(t)$ and $r_Q(t)$ as jointly Gaussian random processes. The Gaussian property also holds for small $N(t)$ if the $\alpha_i(t)$ are Rayleigh distributed (given in (3.36) below) and the $\phi_i(t)$ are uniformly distributed on $[-\pi, \pi]$. This happens, again by the CLT, when the ith multipath component results from a reflection cluster with a large number of nonresolvable multipath components [3]. If the LOS or any of the multipath components has a much larger amplitude than the average multipath component amplitude, then this component dominates the received signal; hence the Gaussian approximation is no longer applicable as the terms in (3.16) are not i.i.d.

3.2.1 Autocorrelation, Cross-Correlation, and Power Spectral Density

We now derive the autocorrelation and cross-correlation of the in-phase and quadrature received signal components $r_I(t)$ and $r_Q(t)$, as well as the received signal $r(t)$. We focus on these correlations over a time period where the transmitter or receiver moves on the order of a few signal wavelengths. As we will see, multipath can cause the received signal to completely decorrelate over a signal wavelength due to changes in phase of the individual multipath components such that they combine either constructively or destructively. Our derivations of these correlations are based on some key assumptions that generally apply to propagation models without a dominant LOS component such that all nonzero terms in (3.16) are approximately i.i.d. We will also determine the power spectral density (PSD) of the received signal under certain assumptions that make its autocorrelation function time-invariant. In this case the PSD is just the Fourier transform of the autocorrelation function.

A transmitter or receiver moving at velocity v transverses a distance equal to a signal wavelength λ over time $\tau = \lambda/v$. We will assume in this section that the LOS and multipath amplitudes $\alpha_i(t)$, delays $\tau_i(t) = d_i(t)/c$, and Doppler frequencies $f_{D_i}(t) = v(t)\cos\theta_i(t)/\lambda$ are changing slowly enough to be considered constant over the time interval τ for which we compute the received signal autocorrelation and cross-correlation. This is generally true for $\tau \approx \lambda/v$ since over time τ or, equivalently, over distance λ the number of signal components $N(t) = N$ doesn't change. Moreover, since the shadowing associated with $\alpha_i(t)$ decorrelates over distance X_c which is generally much bigger than λ, $\alpha_i(t) \approx \alpha_i$ for some constant α_i. Similarly, the distance a given signal component travels, the component's AoA, and the velocity of the transmitter or receiver change little over the time interval τ, and hence we can assume $\tau_i(t) \approx \tau_i$ for some average delay τ_i and $f_{D_i}(t) \approx f_{D_i}$ for some average Doppler frequency f_{D_i}. With a constant Doppler frequency, the Doppler phase shift[4] is $\phi_{D_i}(t) = \int_0^t 2\pi f_{D_i} d\nu = 2\pi f_{D_i} t$ and the phase of the ith multipath component becomes $\phi_i(t) = 2\pi f_c \tau_i - 2\pi f_{D_i} t - \phi_0$.

We now make a key assumption; we assume that for the ith multipath component, the term $2\pi f_c \tau_i - 2\pi f_{D_i} t$ in $\phi_i(t)$ changes rapidly relative to all other phase terms in the expression. This is a reasonable assumption because f_c is large and hence the term $2\pi f_c \tau_i(t)$ can go through a 360° rotation for a small change in multipath delay $\tau_i(t)$. As discussed earlier, this change is typically captured in the Doppler term $2\pi f_{D_i} t$ for $\tau_i(t)$ assumed constant. Under this assumption, $\phi_i(t)$ is uniformly distributed on $[-\pi, \pi]$. Thus

$$\mathbf{E}[r_I(t)] = \mathbf{E}\left[\sum_i \alpha_i \cos\phi_i(t)\right] = \sum_i \mathbf{E}[\alpha_i]\mathbf{E}[\cos\phi_i(t)] = 0, \tag{3.20}$$

[4] We shall assume a Doppler phase shift at $t = 0$ of zero for simplicity, because this phase offset will not affect the analysis.

where the second equality follows from the independence of α_i and $\phi_i(t)$ and the last equality follows from the uniform distribution on $\phi_i(t)$. Similarly we can show that $\mathbf{E}[r_Q(t)] = 0$. Thus, the received signal also has $\mathbf{E}[r(t)] = 0$: it is a zero-mean Gaussian process. If there is a dominant LOS component in the channel then the phase of the received signal is dominated by the phase of the LOS component, which can be determined at the receiver, so the assumption of a random uniform phase no longer holds.

Consider now the autocorrelation $A_{r_I}(t, t+\tau) = \mathbf{E}[r_I(t)r_I(t+\tau)]$ of the in-phase component. Using the independence of α_i and $\phi_i(t)$, the independence of $\phi_i(t)$ and $\phi_j(t)$ $(i \neq j)$, and the uniform distribution of $\phi_i(t)$, we get that

$$\begin{aligned} A_{r_I}(t, t+\tau) &= \mathbf{E}\left[\sum_i \alpha_i \cos\phi_i(t) \sum_j \alpha_j \cos\phi_j(t+\tau)\right] \\ &= \sum_i \sum_j \mathbf{E}[\alpha_i\alpha_j]\mathbf{E}[\cos\phi_i(t)\cos\phi_j(t+\tau)] \\ &= \sum_i \mathbf{E}[\alpha_i^2]\mathbf{E}[\cos\phi_i(t)\cos\phi_i(t+\tau)]. \end{aligned} \tag{3.21}$$

Since $\phi_i(t) = 2\pi f_c\tau_i - 2\pi f_{D_i}t - \phi_0$ and $\phi_i(t+\tau) = 2\pi f_c\tau_i - 2\pi f_{D_i}(t+\tau) - \phi_0$ we have

$$\mathbf{E}[\cos\phi_i(t)\cos\phi_i(t+\tau)] = .5\mathbf{E}[\cos 2\pi f_{D_i}\tau] + .5\mathbf{E}[\cos(4\pi f_c\tau_i - 4\pi f_{D_i}t - 2\pi f_{D_i}\tau - 2\phi_0)]. \tag{3.22}$$

Because $2\pi f_c\tau_i$ changes rapidly relative to all other phase terms and is uniformly distributed modulo 2π, the second expectation term in (3.22) is zero and thus

$$A_{r_I}(t, t+\tau) = .5\sum_i \mathbf{E}[\alpha_i^2]\mathbf{E}[\cos(2\pi f_{D_i}\tau)] = .5\sum_i \mathbf{E}[\alpha_i^2]\mathbf{E}_{\theta_i}\left[\cos\left(\frac{2\pi v\tau}{\lambda}\cos\theta_i\right)\right], \tag{3.23}$$

since $f_{D_i} = v\cos(\theta_i)/\lambda$. The last expectation in (3.23) is with respect to the angle of arrival θ_i since that is the only random parameter in the expression over which the expectation is taken. Observe that $A_{r_I}(t, t+\tau)$ depends only on τ and hence $r_I(t)$ is a wide-sense stationary (WSS) random process. Thus we can write $A_{r_I}(t, t+\tau) = A_{r_I}(\tau)$ as a function of τ only. Moreover, using a similar derivation as (3.23) we can show that the quadrature component is also WSS with the same autocorrelation as $A_{r_I}(\tau)$, so $A_{r_I}(\tau) = A_{r_Q}(\tau)$.

We now consider the cross-correlation between the in-phase and quadrature components. Using the same trigonometric properties as in the derivation of (3.23) and the fact that $2\pi f_c\tau_i$ is uniformly distributed we get

$$A_{r_I,r_Q}(t, t+\tau) = \mathbf{E}[r_I(t)r_Q(t+\tau)] = -.5\sum_i \mathbf{E}[\alpha_i^2]\mathbf{E}_{\theta_i}\left[\sin\left(\frac{2\pi v\tau}{\lambda}\cos\theta_i\right)\right] \tag{3.24}$$

and

$$A_{r_Q,r_I}(t, t+\tau) = \mathbf{E}[r_Q(t)r_I(t+\tau)] = .5\sum_i \mathbf{E}[\alpha_i^2]\mathbf{E}_{\theta_i}\left[\sin\left(\frac{2\pi v\tau}{\lambda}\cos\theta_i\right)\right]. \tag{3.25}$$

We see from (3.24) and (3.25) that these cross-corrections only depend on τ so we can write these as $A_{r_I,r_Q}(t, t+\tau) = A_{r_I,r_Q}(\tau)$ and $A_{r_Q,r_I}(t, t+\tau) = A_{r_Q,r_I}(\tau)$. Moreover, we see from these equations that $A_{r_I,r_Q}(\tau) = -A_{r_Q,r_I}(\tau)$. Finally, setting $\tau = 0$ we see that $A_{r_I,r_Q}(0) = -A_{r_Q,r_I}(0) = 0$. Thus, $r_I(t)$ and $r_Q(t)$ are uncorrelated and, since they are jointly Gaussian processes, this means they are independent.

We now consider the autocorrelation of the received signal

$$r(t) = r_I(t)\cos(2\pi f_c t) - r_Q(t)\sin(2\pi f_c t). \tag{3.26}$$

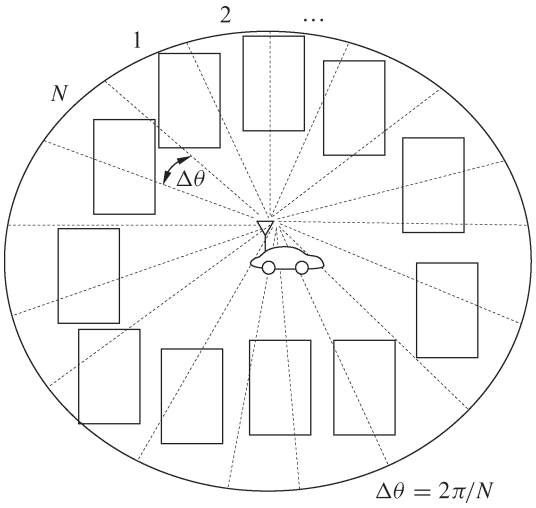


Figure 3.4: Dense scattering environment.

We have

$$\begin{aligned}
A_r(t, t+\tau) &= \mathbf{E}[r(t)r(t+\tau)] \\
&= \mathbf{E}[(r_I(t)\cos(2\pi f_c t) - r_Q(t)\sin(2\pi f_c t))(r_I(t+\tau)\cos(2\pi f_c(t+\tau)) - r_Q(t+\tau)\sin(2\pi f_c(t+\tau))] \\
&= \mathbf{E}[r_I(t)r_I(t+\tau)\cos(2\pi f_c t)\cos(2\pi f_c(t+\tau)) - r_Q(t)r_I(t+\tau)\sin(2\pi f_c t)\cos(2\pi f_c(t+\tau)) \\
&\quad - r_I(t)r_Q(t+\tau))\cos(2\pi f_c t)\sin(2\pi f_c(t+\tau)) + r_Q(t)r_Q(t+\tau)\sin(2\pi f_c t)\sin(2\pi f_c(t+\tau))] \\
&= A_{r_I}(\tau)\cos(2\pi f_c t)\cos(2\pi f_c(t+\tau)) + A_{r_Q}(\tau)\sin(2\pi f_c t)\sin(2\pi f_c(t+\tau)) \\
&\quad - A_{r_Q,r_I}(\tau)\sin(2\pi f_c t)\cos(2\pi f_c(t+\tau)) - A_{r_I,r_Q}(\tau)\cos(2\pi f_c t)\sin(2\pi f_c(t+\tau)) \\
&= A_{r_I}(\tau)\cos(2\pi f_c \tau) - A_{r_I,r_Q}(\tau)\sin(2\pi f_c \tau), \qquad (3.27)
\end{aligned}$$

where the last equality follows from trigonometric product formulas and that fact that $A_{r_I}(\tau) = A_{r_Q}(\tau)$ and $A_{r_I,r_Q}(\tau) = -A_{r_Q,r_I}(\tau)$. Since (3.27) is only a function of the time difference τ, the received signal $r(t)$ is WSS along with its in-phase and quadrature components.

In order to further simplify (3.23) and (3.24), we must make an additional assumption about the propagation environment with respect to the power angle spectrum (PAS), which characterizes the received signal power as a function of its angle of arrival. We will focus on the *uniform scattering environment* introduced by Clarke [4] and further developed by Jakes [5, Chap. 1]. In this model, the channel consists of many scatterers densely packed with respect to their angles of arrival, as shown in Figure 3.4. Note that this is a two-dimensional model, so that the elevation angle of arrival for the signal components is ignored. In the uniform scattering model we assume N (LOS and multipath) signal components with angles of arrival $\theta_i = i\triangle\theta$, $i = 0, \ldots, N-1$, where $\triangle\theta = 2\pi/N$. We also assume that each signal component has the same received power and so $\mathbf{E}[\alpha_i^2] = 2P_r/N$, where P_r is the total received power. Then (3.23) becomes

$$A_{r_I}(\tau) = \frac{P_r}{N}\sum_{i=0}^{N-1}\cos\left(\frac{2\pi v\tau}{\lambda}\cos\left(i\triangle\theta\right)\right). \qquad (3.28)$$

Now making the substitution $N = 2\pi/\triangle\theta$ yields

$$A_{r_I}(\tau) = \frac{P_r}{2\pi} \sum_{i=0}^{N-1} \cos\left(\frac{2\pi v\tau}{\lambda} \cos\left(i\triangle\theta\right)\right) \triangle\theta. \tag{3.29}$$

We now take the limit as the number of scatterers grows to infinity, which corresponds to uniform scattering from all directions. Then $N \to \infty, \triangle\theta \to 0$, and the summation in (3.29) becomes an integral:

$$A_{r_I}(\tau) = \frac{P_r}{2\pi} \int_0^{2\pi} \cos\left(\frac{2\pi v\tau}{\lambda} \cos\theta\right) d\theta = P_r J_0(2\pi f_D \tau), \tag{3.30}$$

where $f_D = 2\pi v/\lambda$ and

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} e^{-jx\cos\theta} d\theta \tag{3.31}$$

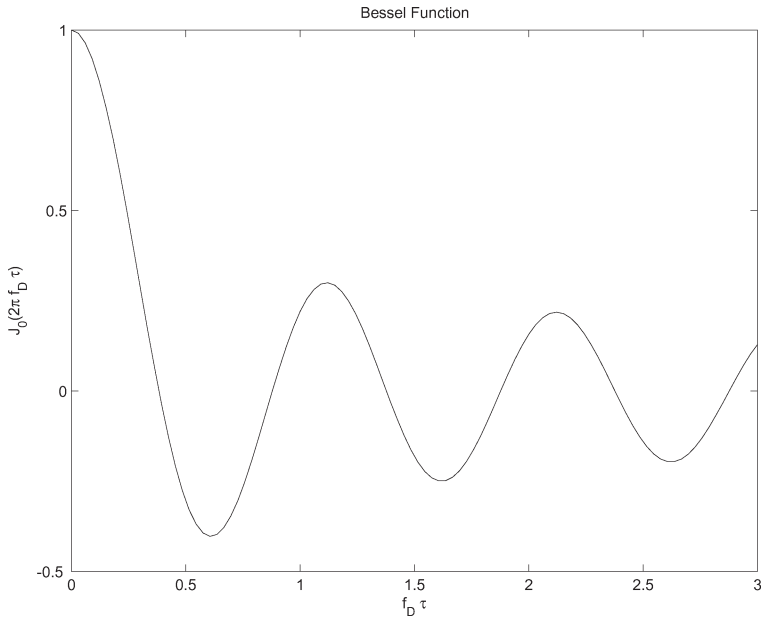


Figure 3.5: Bessel function versus $f_D\tau$.

is the Bessel function of zeroth order.[5] Similarly, for this uniform scattering environment,

$$A_{r_I,r_Q}(\tau) = \frac{P_r}{2\pi} \int \sin\left(\frac{2\pi v\tau}{\lambda} \cos(\theta)\right) d\theta = 0. \tag{3.32}$$

A plot of $J_0(2\pi f_D\tau)$ is shown in Figure 3.5. There are several interesting observations to make from this plot. First we see that the autocorrelation is zero for $f_D\tau \approx .4$ or, equivalently, for $v\tau \approx .4\lambda$. Thus, the signal decorrelates over a distance of approximately one half wavelength under the uniform θ_i assumption. This approximation is commonly used as a rule of thumb to determine many system parameters of interest. For example, we will see in Chapter 7 that independent fading paths obtained from multiple antennas can be combined to remove some of the negative effects of fading. For independent fading the antenna spacing must be $.4\lambda$ based on the analysis here.

[5]Equation (3.30) can also be derived by assuming that $2\pi v\tau\cos\theta_i/\lambda$ in (3.23) and (3.24) is random with θ_i uniformly distributed, and then taking expectations with respect to θ_i. However, based on the underlying physical model, θ_i can be uniformly distributed only in a dense scattering environment. So the derivations are equivalent.

However, combining paths that have a low correlation leads to almost the same gains as combining independent fading paths [6, Chap. 9.6.5]. Another interesting characteristic of this plot is that the signal recorrelates after it becomes uncorrelated. Thus, we cannot assume that the signal remains independent from its initial value at $d = 0$ for separation distances greater than $.4\lambda$. Because of this recorrelation property, a Markov model is not completely accurate for a uniform scattering environment. However, since the recorrelation once the separation distance is greater than a half-wavelength is below .3, this is a reasonable approximation in analyses where such a low correlation has little impact on performance.

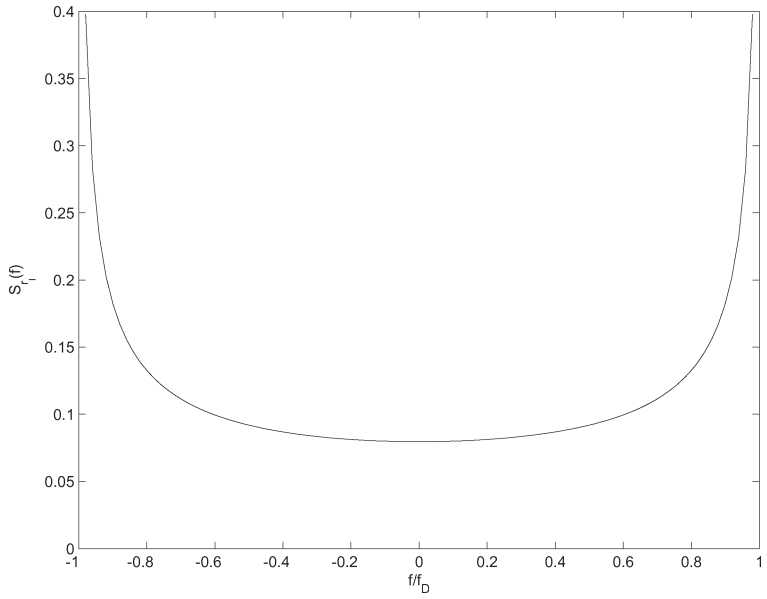


Figure 3.6: In-phase and quadrature PSD: $S_{r_I}(f) = S_{r_Q}(f)$.

The power spectral densities (PSDs) of $r_I(t)$ and $r_Q(t)$ – denoted by $S_{r_I}(f)$ and $S_{r_Q}(f)$, respectively – are obtained by taking the Fourier transform of their respective autocorrelation functions relative to the delay parameter τ. Since these autocorrelation functions are equal, so are the PSDs. Thus

$$S_{r_I}(f) = S_{r_Q}(f) = \mathcal{F}[A_{r_I}(\tau)] = \begin{cases} \frac{2P_r}{\pi f_D} \frac{1}{\sqrt{1-(f/f_D)^2}}, & |f| \le f_D, \\ 0, & \text{otherwise.} \end{cases} \tag{3.33}$$

This PSD is shown in Figure 3.6.

To obtain the PSD of the received signal $r(t)$ under uniform scattering we use (3.27) with $A_{r_I,r_Q}(\tau) = 0$, (3.33), and properties of baseband random processes modulated by a sinusoid (discussed in Appendix B.3) to obtain

$$\begin{aligned} S_r(f) &= \mathcal{F}[A_r(\tau)] = .25[S_{r_I}(f - f_c) + S_{r_I}(f + f_c)] \\ &= \begin{cases} \frac{P_r}{2\pi f_D} \frac{1}{\sqrt{1-(|f-f_c|/f_D)^2}}, & |f - f_c| \le f_D, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{3.34}$$

Note that this PSD integrates to P_r, the total received power.

Since the PSD models the power density associated with multipath components as a function of their Doppler frequency, it can be viewed as the probability density function (pdf) of the random frequency due to Doppler associated with multipath. We see from Figure 3.6 that the PSD $S_{r_I}(f)$ goes to infinity at $f = \pm f_D$ and, consequently, the PSD $S_r(f)$ goes to infinity at $f = \pm f_c \pm f_D$. This will not be true in practice, since the uniform scattering model is just an approximation, but for environments with dense scatterers the PSD will generally be maximized at frequencies close to the maximum Doppler frequency. The intuition for this behavior comes from the nature of the cosine function and the fact that (under our assumptions) the PSD corresponds to the pdf of the random Doppler frequency $f_D(\theta)$.

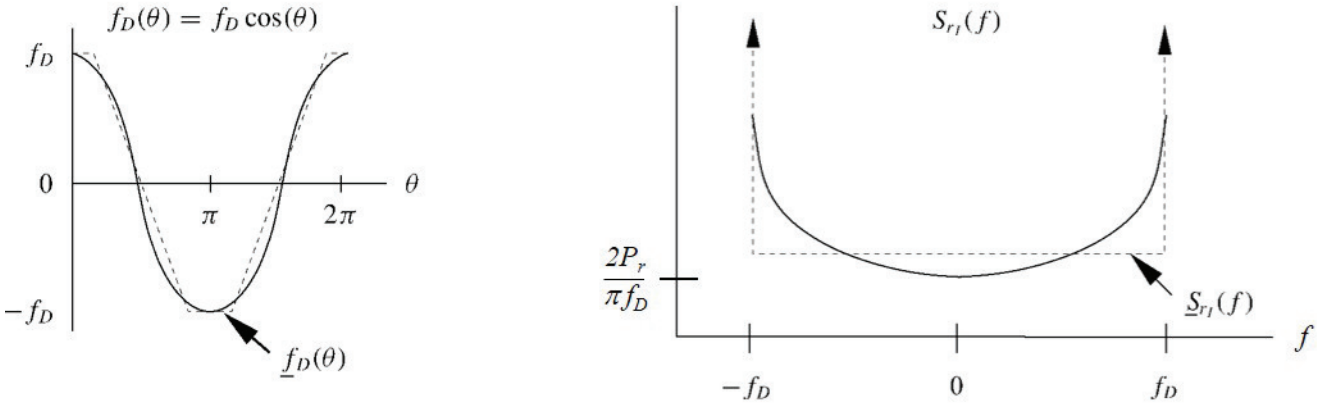


Figure 3.7: Cosine and PSD approximation by straight line segments.

To see this, note that the uniform scattering assumption is based on many scattered paths arriving uniformly from all angles with the same average power. Thus, θ for a randomly selected path can be regarded as a uniform random variable on $[0, 2\pi]$. The pdf $p_{f_D(\theta)}(f)$ of the random Doppler frequency $f_D(\theta)$ can then be obtained from the pdf of θ. By definition, $p_{f_D(\theta)}(f)$ is proportional to the density of scatterers at Doppler frequency f. Hence, $S_{r_I}(f)$ is also proportional to this density, and we can characterize the PSD from the pdf $p_{f_D(\theta)}(f)$. For this characterization, in Figure 3.7 we plot $f_D(\theta) = f_D \cos(\theta) = \frac{v}{\lambda}\cos(\theta)$ along with a (dashed) straight-line segment approximation $\underline{f}_D(\theta)$ to $f_D(\theta)$. On the right in this figure we plot the PSD $S_{r_I}(f)$ along with a dashed straight-line segment approximation to it, $\underline{S}_{r_I}(f)$, which corresponds to the Doppler approximation $\underline{f}_D(\theta)$. We see that $\cos(\theta) \approx \pm 1$ for a relatively large range of θ-values. Thus, multipath components with angles of arrival in this range of values have Doppler frequency $f_D(\theta) \approx \pm f_D$, so the power associated with all of these multipath components will add together in the PSD at $f \approx f_D$. This is shown in our approximation by the fact that the segments where $\underline{f}_D(\theta) = \pm f_D$ on the left lead to delta functions at $\pm f_D$ in the PSD approximation $\underline{S}_{r_I}(f)$ on the right. The segments where $\underline{f}_D(\theta)$ has uniform slope on the left lead to the flat part of $\underline{S}_{r_I}(f)$ on the right, since there is one multipath component contributing power at each angular increment. This explains the shape of $S_{r_I}(f)$ under uniform scattering. Formulas for the autocorrelation and PSD in nonuniform scattering – corresponding to more typical microcell and indoor environments – can be found in [5, Chap. 1] and [7, Chap. 2]. Three-dimensional channel models that incorporate elevation angle are described in [8] and the references therein.

The PSD is useful in constructing simulations for the fading process. A common method for simulating the envelope of a narrowband fading process is to pass two independent white Gaussian noise sources with PSD $N_0/2$ through lowpass filters with a frequency response $H(f)$ that satisfies

$$S_{r_I}(f) = S_{r_Q}(f) = \frac{N_0}{2}|H(f)|^2. \tag{3.35}$$

The filter outputs then correspond, respectively, to the in-phase and quadrature components of the narrowband fading process with PSDs $S_{r_I}(f)$ and $S_{r_Q}(f)$. A similar procedure using discrete filters can be used to generate discrete fading processes. Most communication simulation packages (e.g. the Matlab communications toolbox) have standard modules that simulate narrowband fading based on this method. More details on this simulation method, as well as alternative methods, can be found in [1, 7, 9, 10].

We have now completed our model for the three characteristics of power versus distance exhibited in narrowband wireless channels. These characteristics are illustrated in Figure 3.8, adding narrowband fading to the single-slope path loss and log-normal shadowing models developed in Chapter 2. In this figure we see the signal power due to path loss decreasing at a slope of -10γ relative to $\log_{10}(d/d_r)$ for γ the path-loss exponent and d_r the reference distance at which the path loss equals the constant K. The more rapid variations due to shadowing change on the order of the decorrelation distance X_c, and the very rapid variations due to multipath fading change on the order of half the signal wavelength. If we blow up a small segment of this figure over distances where path loss and shadowing are constant we obtain Figure 3.9, which plots the dB value of P_r/P_t versus linear distance $d = vt$ (not log distance). In this figure the average value of P_r/P_t is normalized to 0 dB. A mobile receiver traveling at fixed velocity v would experience stationary and ergodic received power variations over time as illustrated in this figure.

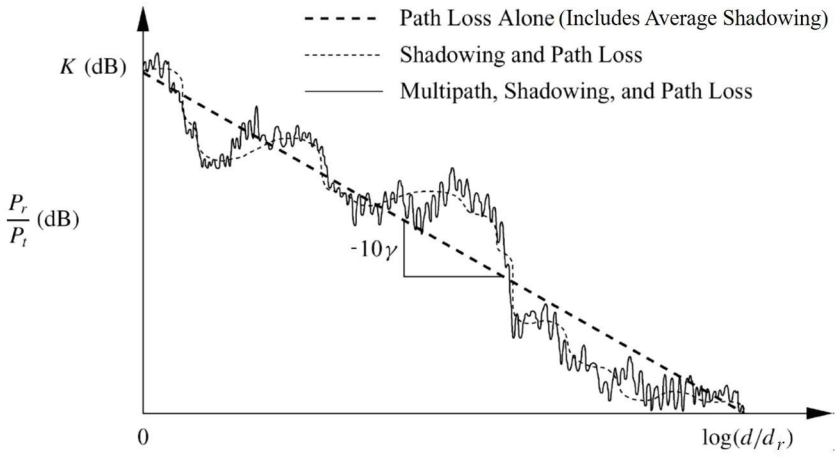


Figure 3.8: Combined single-slope path loss, log-normal shadowing, and narrowband fading.

3.2.2 Envelope and Power Distributions for Narrowband Fading

We now consider the distribution of the envelope and power for the narrowband received signal $r(t) = r_I(t)\cos(2\pi f_c t) - r_Q(t)\sin(2\pi f_c t)$ when the path loss and shadowing for each multipath component is roughly constant, i.e. we assume that $\alpha_i(t) = \alpha_i$ in (3.17) and (3.18). Thus, random variations of the received signal envelope and power are due only to the constructive and destructive combining of the different multipath components. These signal envelope and power distributions depend on whether there is one or more dominant multipath components in the received signal or, alternatively, if all the multipath components have roughly the same received power. As will be derived below, without a dominant multipath component and assuming a uniform angle of arrival distribution for all multipath components, the signal envelope follows a *Rayleigh* distribution. When there is a single dominant multipath component among a large number of multipath components with similar power and uniform angle of arrival, for example due to a strong LOS component, the signal envelope distribution is *Rician*. There are different mathematical and empirical models for channels with two or more dominant multipath components or other dif-

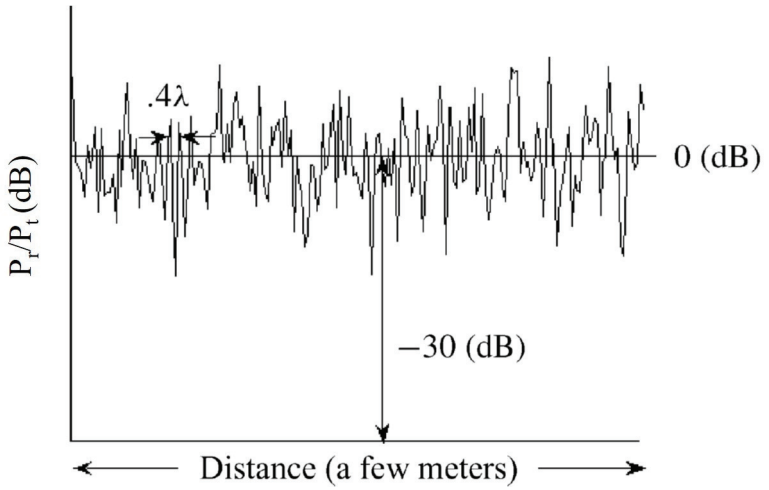


Figure 3.9: Narrowband fading.

ferences with the models that lead to Rayleigh and Rician fading.

RAYLEIGH FADING

We derived in Section 3.2.1 that for $\alpha_i(t) = \alpha_i$ and $\phi_i(t)$ uniformly distributed, r_I and r_Q are both zero-mean Gaussian random variables, meaning that the received signal does not have a dominant multipath component. It can be shown that, for any two Gaussian random variables X and Y, both with mean zero and equal variance σ^2, $Z = \sqrt{X^2 + Y^2}$ is Rayleigh distributed and Z^2 is exponentially distributed. If we assume a variance of $\sigma^2 = .5\sum_i \mathbf{E}[\alpha_i^2]$ for both in-phase and quadrature components, then the signal envelope $z(t) = |r(t)| = \sqrt{r_I^2(t) + r_Q^2(t)}$ is Rayleigh distributed with distribution

$$p_Z(z) = \frac{2z}{P_r}\exp\left[-\frac{z^2}{P_r}\right] = \frac{z}{\sigma^2}\exp\left[-\frac{z^2}{2\sigma^2}\right], \quad z \geq 0, \tag{3.36}$$

where $P_r = 2\sigma^2$ is the power of $z(t)$ which equals the power of $r(t)$.

We obtain the power distribution by making the change of variables $z^2(t) = |r(t)|^2$ in (3.36) to obtain

$$p_{Z^2}(x) = \frac{1}{P_r}e^{-x/P_r} = \frac{1}{2\sigma^2}e^{-x/2\sigma^2}, \quad x \geq 0, \tag{3.37}$$

i.e. the power in $r(t)$ is exponentially distributed with mean $P_r = 2\sigma^2$. Thus, $r(t)$ has a Rayleigh-distributed amplitude and exponentially-distributed power with mean $P_r = 2\sigma^2$. The equivalent lowpass signal for $r(t)$ is given by $r_{LP}(t) = r_I(t) + jr_Q(t)$, which has the same power as $z(t)$ and phase $\theta = \arctan\,(r_Q(t)/r_I(t))$. For $r_I(t)$ and $r_Q(t)$ uncorrelated Gaussian random variables we can show that θ is uniformly distributed and independent of $|r_{LP}|$.

Example 3.2: Consider a channel with Rayleigh fading and average received power $P_r = 20$ dBm. Find the probability that the received power is below 10 dBm.

Solution: We have $P_r = 20\ \mathrm{dBm} = 100\ \mathrm{mW}$. We want to find the probability that $Z^2 < 10\ \mathrm{dBm} = 10\ \mathrm{mW}$. Thus

$$p(Z^2 < 10) = \int_0^{10} \frac{1}{100} e^{-x/100} dx = .095. \tag{3.38}$$

RICIAN FADING

If the channel has a dominant LOS or other multipath component with a much larger signal power than the other multipath components, then $r_I(t)$ and $r_Q(t)$ are not zero-mean random processes. That is because the angle of arrival for the dominant multipath component is not random and, in particular, it can no longer be modeled as uniformly distributed. In this case the average signal power equals that of the dominant multipath component, with small fluctuations about this average due to constructive and destructive addition of the other multipath components. Assuming the dominant multipath component is its LOS component, the in-phase and quadrature received signal components can be written as

$$r_I = \alpha_0 \cos\phi_0 + \sum_{i=1}^{N-1} \alpha_i \cos\phi_i \tag{3.39}$$

and

$$r_Q = \alpha_0 \sin\phi_0 + \sum_{i=1}^{N-1} \alpha_i \sin\phi_i(t). \tag{3.40}$$

When N is large and ϕ_i is uniformly distributed for $i \neq 0$, the second term in these expressions corresponds to the Rayleigh fading model above, i.e. $\sum_{i=1}^{N-1} \alpha_i \cos\phi_i$ and $\sum_{i=1}^{N-1} \alpha_i \sin\phi_i$ correspond to zero-mean Gaussian random variables. Hence the received signal equals the superposition of a complex Gaussian component and a higher-power single multipath component with amplitude α_0 and phase ϕ_0. The signal envelope in this case can be shown to have a Rician distribution [11] given by

$$p_Z(z) = \frac{z}{\sigma^2} \exp\left[\frac{-(z^2+s^2)}{2\sigma^2}\right] I_0\left(\frac{zs}{\sigma^2}\right), \quad z \geq 0, \tag{3.41}$$

where $2\sigma^2 = \sum_{i,i\neq 0} \mathbf{E}[\alpha_i^2]$ is the average power in the non-LOS multipath components and $s^2 = \alpha_0^2$ is the power in the LOS component. The function I_0 is the modified Bessel function of zeroth order. The average received power in the Rician fading is given by

$$P_r = \int_0^{\infty} z^2 p_Z(z) dz = s^2 + 2\sigma^2. \tag{3.42}$$

The Rician distribution is often described in terms of a fading parameter K, defined by

$$K = \frac{s^2}{2\sigma^2}. \tag{3.43}$$

Thus, K is the ratio of the power in the LOS component to the power in the other (non-LOS) multipath components, which is typically a random variable. For $K = 0$ we have Rayleigh fading and for $K = \infty$ we have no fading (i.e., a channel with no multipath and only a LOS component). The fading parameter K is therefore a measure of the severity of the fading: a small K implies severe fading, a large K implies relatively mild fading. Making the substitutions $s^2 = KP_r/(K+1)$ and $2\sigma^2 = P_r/(K+1)$, we can write the Rician distribution in terms of K and P_r as

$$p_Z(z) = \frac{2z(K+1)}{P_r} \exp\left[-K - \frac{(K+1)z^2}{P_r}\right] I_0\left(2z\sqrt{\frac{K(K+1)}{P_r}}\right), \quad z \geq 0. \tag{3.44}$$

NAKAGAMI FADING

Both the Rayleigh and Rician distributions for the received signal envelope can be obtained by using mathematics to capture the underlying physical properties of the channel models [3, 11]. However, some experimental data does not fit well into either of these distributions. Thus, a more general fading distribution for the received signal envelope was developed whose parameters can be adjusted to fit a variety of empirical measurements. This distribution is called the Nakagami fading distribution and is given by

$$p_Z(z) = \frac{2m^m z^{2m-1}}{\Gamma(m) P_r^m} \exp\left[\frac{-mz^2}{P_r}\right], \quad m \geq .5, \tag{3.45}$$

where P_r is the average received power and $\Gamma(\cdot)$ is the Gamma function. The Nakagami distribution is parameterized by P_r and the fading parameter m. For $m = 1$ the distribution in (3.45) reduces to Rayleigh fading. For $m = (K+1)^2/(2K+1)$ the distribution in (3.45) is approximately Rician fading with parameter K. For $m = \infty$ there is no fading: $Z = \sqrt{P_r}$ is a constant. Thus, the Nakagami distribution can model both Rayleigh and Rician distributions as well as more general ones. Note that some empirical measurements support values of the m-parameter less than unity, in which case the Nakagami fading causes more severe performance degradation than Rayleigh fading. The power distribution for Nakagami fading, obtained by a change of variables, is given by

$$p_{Z^2}(x) = \left(\frac{m}{P_r}\right)^m \frac{x^{m-1}}{\Gamma(m)} \exp\left[\frac{-mx}{P_r}\right]. \tag{3.46}$$

The Nakagami fading model has been generalized with respect to the means and variances of its in-phase and quadrature components. In particular, the Hoyt (or Nakagami-q) distribution models narrowband fading of the received signal envelope assuming instantaneous values of the in-phase and quadrature components are zero-mean Gaussian random variables with unequal variances. The Beckman distribution generalizes the Hoyt model such that the instantaneous in-phase and quadrature components are Gaussian random variables with unequal means and unequal variances. The pdf formulas for the Hoyt and Beckman distributions can be found in [6].

MULTI-RAY STATISTICAL FADING MODEL

The Rician fading model corresponding to the superposition of one dominant multipath component and a large number of additional multipath components of lower power can be generalized to include multiple dominant multipath components, as introduced in [12]. Assuming all the multipath components have i.i.d. uniformly-distributed phase and are numbered based on their relative average power, with N dominant multipath components and M additional lower-power components, the in-phase and quadrature signal components in this model are given by

$$r_I = \sum_{i=0}^{N-1} \alpha_i \cos\phi_i + \sum_{i=N}^{N+M-1} \alpha_i \cos\phi_i \tag{3.47}$$

and

$$r_Q = \sum_{i=0}^{N-1} \alpha_i \sin\phi_i + \sum_{i=N}^{N+M-1} \alpha_i \sin\phi_i. \tag{3.48}$$

Similar to the Rician case, when M is large the second summations in (3.47) and (3.50) converge to zero-mean Gauss-distributed random variables. When both M and N are large the model reduces to the Rayleigh fading

model since both summations in these expressions are equivalent in that they both converge to zero-mean Gauss-distributed random variables. Hence, the multi-ray statistical fading model is used primarily to model channels with a small number N of dominant multipath components and a large number M of additional lower-power components.

A special case of the multi-ray statistical model considers just two dominant multipath components, i.e. $N = 2$. This leads to the following in-phase and quadrature signal components:

$$r_I = \alpha_0 \cos\phi_0 + \alpha_1 \cos\phi_1 + \sum_{i=2}^{M+1} \alpha_i \cos\phi_i \tag{3.49}$$

and

$$r_Q = \alpha_0 \sin\phi_0 + \alpha_1 \sin\phi_1 + \sum_{i=2}^{M+1} \alpha_i \sin\phi_i. \tag{3.50}$$

When α_0 and α_1 are deterministic, this model is referred to as the *Two-wave with Diffuse Power* (TWDP) model [12]. The TWDP model is a generalization of the deterministic two-ray multipath model described in Section 2.4, whereby the two multipath components in that model are combined with a large number of lower-power multipath components, and all components are assumed to have i.i.d. uniform phase. This model is a good approximation for channels with a LOS and ground reflection along with multipath components that undergo significant shadowing, path loss, and scattering. Although there is no closed-form expression for the envelope or power distribution of the TWDF model, close approximations are derived in [12]. When α_0 and α_1 follow a Nakagami distribution, the model is referred to as the *Fluctuating Two-Ray* (FTR) fading model. Closed-form expressions for the power distribution of the FTR fading model are derived in [13]. This model has been shown to provide a good approximation for some millimeter wave channels where the dominant multipath components experience fading due to rough reflections, precipitation, and other propagation phenomena.

3.2.3 Level Crossing Rate and Average Fade Duration

In this section we derive the average fade duration of a narrowband fading signal, which equals the average length of time the signal is continuously below a given value. To compute the average fade duration, we must first compute the signal's level crossing rate. The level crossing rate is a characteristic of any random process. It is defined as the expected rate, in crossings per second, that the process crosses a particular value or level. The rate can be based on crossing the level in either the upward or downward direction. In order to compute the average fade duration, we consider the level crossing rate in the downward direction, since the signal is then below the desired value. We thus define the envelope level crossing rate L_Z of signal envelope $z(t) = |r(t)|$ as the expected rate, in crossings per second, at which the signal envelope crosses the level Z in the downward direction. Obtaining L_Z requires the joint distribution $p(z, \dot{z})$ of the signal envelope $z = |r|$ and its derivative with respect to time, $\dot{z}$. We now derive L_Z based on this joint distribution.

Consider the fading process shown in Figure 3.10. The expected amount of time that the signal envelope spends in the interval $(Z + dz, Z)$ with envelope slope in the range $(\dot{z} + d\dot{z}, \dot{z})$ over time duration dt is $A = p(Z, \dot{z})\, dz\, d\dot{z}\, dt$. The time required to cross from Z to $Z + dz$ once for a given envelope slope $\dot{z}$ is $B = dz/\dot{z}$. The ratio $A/B = \dot{z}p(Z, \dot{z})\, d\dot{z}\, dt$ is the expected number of crossings of the envelope z within the interval $(Z + dz, Z)$ for a given envelope slope $\dot{z}$ over time duration dt. The expected number of crossings of the envelope level Z for slopes between $\dot{z} + d\dot{z}$ and $\dot{z}$ in a time interval $[0, T]$ in the downward direction is thus

$$\int_0^T \dot{z}p(Z, \dot{z}) d\dot{z}\, dt = \dot{z}p(Z, \dot{z}) d\dot{z} T. \tag{3.51}$$

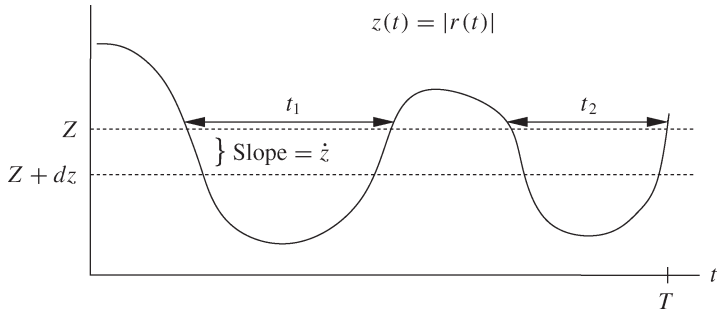


Figure 3.10: Level crossing rate and fade duration for fading process.

Hence the expected number of crossings of the envelope level Z with negative slope over the interval $[0, T]$ is

$$N_Z = T \int_{-\infty}^{0} \dot{z} p(Z, \dot{z}) d\dot{z}. \tag{3.52}$$

Finally, the expected number of crossings of the envelope level Z per second – that is, the level crossing rate – is

$$L_Z = \frac{N_Z}{T} = \int_{-\infty}^{0} \dot{z} p(Z, \dot{z}) d\dot{z}. \tag{3.53}$$

Note that this derivation of level crossing rate is a general result that applies for any random process.

The joint pdf of z and $\dot{z}$ for Rician fading was derived in [11] and can also be found in [7, Chap. 2.1]. The level crossing rate for Rician fading is then obtained by using this pdf in (3.53), yielding

$$L_Z = \sqrt{2\pi(K+1)} f_D \rho e^{-K-(K+1)\rho^2} I_0(2\rho\sqrt{K(K+1)}), \tag{3.54}$$

where $\rho = Z/\sqrt{\bar{P}_r}$. For Rayleigh fading $(K = 0)$ the level crossing rate simplifies to

$$L_Z = \sqrt{2\pi} f_D \rho e^{-\rho^2}. \tag{3.55}$$

We define the average signal fade duration as the average time that the signal envelope stays below a given target level Z. This target level is often obtained from the signal amplitude or power level required for a given performance metric such as bit error rate. If the signal amplitude or power falls below its target then we say the system is in outage. Let t_i denote the duration of the ith fade below level Z over a time interval $[0, T]$, as illustrated in Figure 3.10. Thus t_i equals the length of time that the signal envelope stays below Z on its ith crossing. Since $z(t)$ is stationary and ergodic, for T sufficiently large we have

$$p(z(t) < Z) = \frac{1}{T} \sum_i t_i. \tag{3.56}$$

Thus, for T sufficiently large, the average fade duration is

$$\bar{t}_Z = \frac{1}{TL_Z} \sum_{i=1}^{L_Z T} t_i \approx \frac{p(z(t) < Z)}{L_Z}. \tag{3.57}$$

Using the Rayleigh distribution for $p(z(t) < Z)$ then yields

$$\bar{t}_Z = \frac{e^{\rho^2} - 1}{\rho f_D \sqrt{2\pi}} \tag{3.58}$$

with $\rho = Z/\sqrt{\bar{P}_r}$. Note that (3.58) is the average fade duration for the signal envelope (amplitude) level with Z the target amplitude and $\sqrt{\bar{P}_r}$ the average envelope level. By a change of variables it is easily shown that (3.58) also yields the average fade duration for the signal power level with $\rho = \sqrt{P_0/\bar{P}_r}$, where P_0 is the target power level and $\bar{P}_r$ is the average power level. The average fade duration (3.58) decreases with the Doppler frequency f_D, since as a channel changes more quickly it remains below a given fade level for a shorter period of time. The average fade duration also generally increases with ρ for $\rho \gg 1$. That is because the signal is more likely to be below the target as the target level increases relative to the average. The average fade duration for Rician fading is more difficult to compute; it can be found in [7, Chap. 1.4].

The average fade duration indicates the number of bits or symbols affected by a deep fade. Specifically, consider an uncoded system with bit time T_b. Suppose the probability of bit error is high when $z < Z$. In this case, if $T_b \approx \bar{t}_Z$ then the system will likely experience single error events, where bits that are received in error have the previous and subsequent bits received correctly (since $z > Z$ for these bits). On the other hand, if $T_b \ll \bar{t}_Z$ then many subsequent bits are received with $z < Z$, so large bursts of errors are likely. Finally, if $T_b \gg \bar{t}_Z$ then, since the fading is integrated over a bit time in the demodulator, the fading gets averaged out and so can be neglected. These issues will be explored in more detail in Chapter 8, where we consider coding and interleaving.

Example 3.3: Consider a voice system with acceptable BER when the received signal power is at or above half its average value. If the BER is below its acceptable level for more than 120 ms, users will turn off their phone. Find the range of Doppler values in a Rayleigh fading channel such that the average time duration when users have unacceptable voice quality is less than $t = 60$ ms.

Solution: The target received signal value is half the average, so $P_0 = .5\bar{P}_r$ and thus $\rho = \sqrt{.5}$. We require

$$\bar{t}_Z = \frac{e^{.5} - 1}{f_D\sqrt{\pi}} \leq t = .060 \tag{3.59}$$

and thus $f_D \geq (e^{.5} - 1)/(.060\sqrt{\pi}) = 6.1$ Hz.

3.2.4 Block-Fading and Finite-State Markov Fading

The complex mathematical characterization of the narrowband fading model described in the previous sections can be difficult to incorporate into wireless performance analysis. Therefore, simpler models that capture the main features of narrowband fading channels are often used for these analytical calculations. This section describes two such models: block-fading and finite-state Markov fading channels.

In a block-fading channel, the channel SNR γ is assumed constant over some period of time T called the *block time*. In this model we denote $\gamma(i)$ as the channel SNR over the ith block. It is typically assumed that $\gamma(i)$ is i.i.d. between blocks although a more complex model capturing correlation of the channel gain across blocks can also be used. Often T is set equal to an integer number of modulation symbol times T_s: $T = NT_s$ where NT_s is approximately equal to the channel coherence time defined in Section 3.3.4 below. A rapidly-fading channel might be approximated by setting $N = 1$ such that the fading is constant over each symbol time and changes to a different i.i.d. value between subsequent symbols. A variation on the block-fading model allows for continuous variation of the channel gain within a block, typically using the model of Section 3.2.1, with an abrupt transition to an independent channel gain at each block time. This model is often used for systems where blocks of symbols are sent intermittently and hence each block experiences a completely different propagation environment.

The finite-state Markov channel (FSMC) is a special case of block-fading where a finite set of channel SNR values is assumed with a Markov transition probability of the fading between blocks. Specifically in the FSMC

the channel SNR is assumed to belong to a finite set of values, called channel states, and to follow a discrete-time Markov process with time discretized to the block time T. The channel varies over its possible states at each interval T according to a set of Markov transition probabilities. FSMCs have been used to approximate both mathematical and experimental fading models, including satellite channels [14], indoor channels [15], Rayleigh fading channels [16, 17], Rician fading channels [18], and Nakagami-m fading channels [19]. They have also been used for system design and system performance analysis [17, 20]. First-order FSMC models are deficient when the fading recorrelation affects performance analysis, in which case higher-order models are used. The FSMC models for fading typically model amplitude variations only, although there has been some work on FSMC models for phase in fading [21] or phase-noisy channels [22].

A detailed FSMC model for Rayleigh fading was developed in [16]. In this model the time-varying SNR γ associated with the Rayleigh fading lies in the range $0 \leq \gamma < \infty$. The FSMC model discretizes this fading range into regions so that the jth region R_j is defined as $R_j = \{\gamma \ : \ A_j \leq \gamma < A_{j+1}\}$, where the region boundaries $\{A_j\}$ and the total number of fade regions are parameters of the model. This model assumes that γ stays within the same region over time interval T and can only transition to the same region or adjacent regions at time $T+1$. Thus, given that the channel is in state R_j at time T, at the next time interval the channel can only transition to R_{j-1}, R_j, or R_{j+1} – a reasonable assumption when $f_D T$ is small. Under this assumption, the transition probabilities between regions are derived in [16] as

$$p_{j,j+1} = \frac{L_{j+1}T}{\pi_j}, \quad p_{j,j-1} = \frac{L_j T}{\pi_j}, \quad p_{j,j} = 1 - p_{j,j+1} - p_{j,j-1}, \tag{3.60}$$

where L_j is the level crossing rate at A_j and π_j is the steady-state distribution corresponding to the jth region: $\pi_j = p(\gamma \in R_j) = p(A_j \leq \gamma < A_{j+1})$.

3.3 Wideband Fading Model

Suppose the transmitted signal is not narrowband relative to the inverse delay spread. Since the signal bandwidth $B \approx 1/T$ for T the signal duration, the received signal will have duration $T + T_m \gg T$ for T_m is the delay spread. Thus, the duration of the received signal will be significantly increased, and we refer to the channel in this case as a *dispersive* channel. This phenomenon is illustrated in Figure 3.11. In the figure, a pulse of width T is transmitted over a multipath channel. As discussed in Chapter 5, linear modulation consists of a train of pulses where each pulse carries information in its amplitude and/or phase corresponding to a data bit or symbol.[6] If the multipath delay spread $T_m \ll T$ then the multipath components are received roughly on top of one another, as shown in the upper right of the figure. Hence the channel is nondispersive. The resulting constructive and destructive interference causes narrowband fading of the pulse, but there is little time spreading of the pulse and therefore little interference with a subsequently transmitted pulse. On the other hand, if the multipath delay spread $T_m \gg T$, then each of the different multipath components can be resolved, as shown in the lower right of the figure. However, these multipath components interfere with subsequently transmitted pulses (dashed pulses in the figure). This effect is called *intersymbol interference* (ISI).

There are several techniques to mitigate the distortion due to multipath delay spread, including equalization, multicarrier modulation, and spread spectrum, which are discussed in Chapters 11–13. Mitigating ISI is not necessary if $T \gg T_m$. However, most wireless systems require data rates $R \approx 1/T$ to be much greater than $1/T_m$ for typical delay spreads of both indoor and outdoor channels. Multicarrier modulation and spread spectrum actually change the characteristics of the transmitted signal to mostly avoid intersymbol interference; however, they still experience multipath distortion due to frequency-selective fading, which is described in Section 3.3.3.

[6] Linear modulation typically uses nonsquare pulse shapes for bandwidth efficiency, as discussed in Section 5.4.

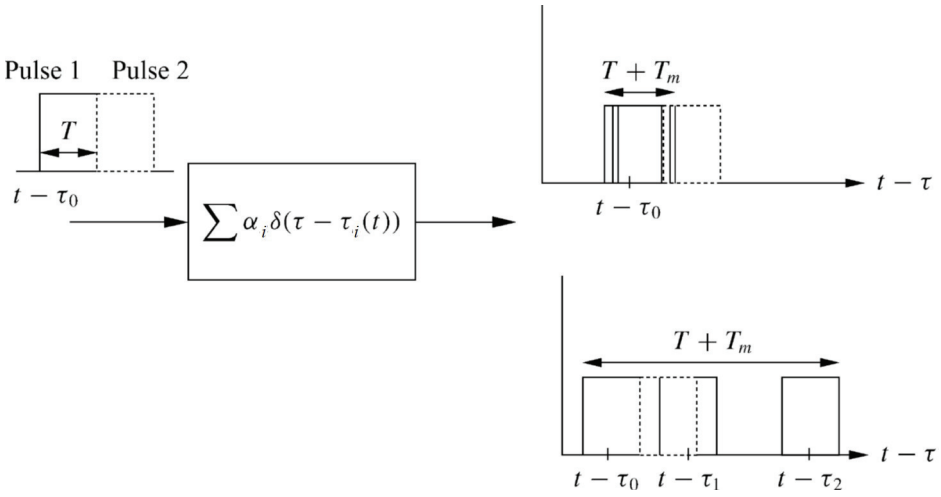


Figure 3.11: Multipath resolution.

The difference between wideband and narrowband fading models is that, as the baseband signal bandwidth B_u increases so that $T_m \approx B_u^{-1}$, the approximation $u(t-\tau_i(t)) \approx u(t)$ is no longer valid for all multipath component delays $\tau_i(t)$. Thus, the received signal is a sum of all copies of the original signal, where each copy is delayed in time by $\tau_i(t)$ and shifted in phase by $\phi_i(t)$. The signal copies will combine destructively when their phase terms differ significantly and will distort the direct path signal when $u(t-\tau_i(t))$ differs from $u(t)$.

Although the approximation in (3.14) no longer applies when the signal bandwidth is large relative to the inverse of the multipath delay spread, if the number of multipath components is large and the phase of each component is uniformly distributed then the received signal will still be a zero-mean complex Gaussian process with a Rayleigh-distributed envelope. However, wideband fading differs from narrowband fading in terms of the resolution of the different multipath components. Specifically, for narrowband signals, the multipath components are not resolvable if their delay difference is less than the inverse of the signal bandwidth, so the multipath components characterized in (3.6) combine at the receiver to yield the original transmitted signal with amplitude and phase characterized by random processes. These random processes are in turn characterized by their autocorrelation (or PSD) and instantaneous distributions, as discussed in Section 3.2. However, with wideband signals, the received signal experiences distortion due to the delay spread of the different multipath components, so the received signal can no longer be characterized by just the amplitude and phase random processes. The effect of multipath on wideband signals must therefore take into account both the multipath delay spread and the time variations associated with the channel.

3.3.1 Autocorrelation and Scattering Function

The starting point for characterizing wideband channels is the equivalent lowpass time-varying channel impulse response $c(\tau,\ t)$. Let us first assume that $c(\tau,\ t)$ is a continuous[7] deterministic function of τ and t. Recall that $c(\tau,\ t)$ represents the response of the channel at time t to an impulse at time $t-\tau$. Hence, time variations of the impulse response are captured by the parameter t. We can take the Fourier transform of $c(\tau,\ t)$ with respect to t as

$$C(\tau,\ \rho) = \int_{-\infty}^{\infty} c(\tau,\ t) e^{-j2\pi\rho t} dt. \tag{3.61}$$

[7]The wideband channel characterizations in this section can also be made for discrete-time channels (discrete with respect to τ) by changing integrals to sums and Fourier transforms to discrete Fourier transforms.

We call $C(\tau,\ \rho)$ the *deterministic scattering function* of the equivalent lowpass channel impulse response $c(\tau,\ t)$. Recall that time variations of the channel due to movement of the transmitter or receiver result in a Doppler frequency shift f_D. Since $C(\tau,\ \rho)$ is the Fourier transform of $c(\tau,\ t)$ with respect to the time variation parameter t, the deterministic scattering function $C(\tau,\ \rho)$ captures the Doppler characteristics of the channel via the frequency parameter ρ. The derivation of this relationship for random channels is given in Section 3.3.4

In general, the time-varying channel impulse response $c(\tau,\ t)$ given by (3.6) is random instead of deterministic because of the random amplitudes, phases, and delays of the random number of multipath components. In this case we must characterize it statistically or via measurements. As long as the number of multipath components is large and the LOS signal component does not dominate the received signal, we can invoke the CLT to assume that $c(\tau,\ t)$ is a zero-mean complex Gaussian process and hence that its statistical characterization is fully known from the mean, autocorrelation, and cross-correlation of its in-phase and quadrature components. As in the narrowband case, we assume that the phase of each multipath component is uniformly distributed. Thus, the in-phase and quadrature components of $c(\tau,\ t)$ are independent Gaussian processes with the same autocorrelation, a mean of zero, and a cross-correlation of zero. The same statistics hold for the in-phase and quadrature components if the channel contains only a small number of multipath rays – as long as each ray has a Rayleigh-distributed amplitude and uniform phase. Note that this model does not hold when the channel has a dominant LOS component.

The statistical characterization of $c(\tau,\ t)$ is thus determined by its *autocorrelation function*, defined as

$$A_c(\tau_1,\ \tau_2; t,\ t+\triangle t) = \mathbf{E}[c^*(\tau_1; t)c(\tau_2; t+\triangle t)]. \tag{3.62}$$

Most time-varying channels associated with deployed wireless systems are WSS, so that the joint statistics of a channel measured at two different times t and $t+\triangle t$ depends only on the time difference $\triangle t$. For WSS channels, the autocorrelation of the corresponding bandpass channel $h(\tau,\ t) = \mathrm{Re}\{c(\tau,\ t)e^{j2\pi f_C t}\}$ can be obtained from $A_c(\tau_1,\ \tau_2; t,\ t+\triangle t)$ as[8] $A_h(\tau_1,\ \tau_2; t,\ t+\triangle t) = .5\,\mathrm{Re}\{A_c(\tau_1,\ \tau_2; t,\ t+\triangle t)e^{j2\pi f_c \triangle t}\}$ [23]. We will assume that our channel model is WSS, in which case the autocorrelation becomes independent of t:

$$A_c(\tau_1,\ \tau_2; \triangle t) = \mathbf{E}[c^*(\tau_1; t)c(\tau_2; t+\triangle t)]. \tag{3.63}$$

Moreover, in real environments the channel response associated with a given multipath component of delay τ_1 is uncorrelated with the response associated with a multipath component at a different delay $\tau_2 \neq \tau_1$, since the two components are caused by different scatterers. We say that such a channel has uncorrelated scattering (US). We denote channels that are WSS with US as WSSUS channels. The WSSUS channel model was first introduced by Bello in his landmark paper [23], where he also developed two-dimensional transform relationships associated with this autocorrelation. These relationships will be discussed in Section 3.3.5. Incorporating the US property into (3.63) yields

$$\mathbf{E}[c^*(\tau_1; t)c(\tau_2; t+\triangle t)] = A_c(\tau_1; \triangle t)\delta[\tau_1 - \tau_2]. \tag{3.64}$$

Now setting $\tau_1 = \tau_2 = \tau$ we get

$$\mathbf{E}[c^*(\tau; t)c(\tau; t+\triangle t)] = A_c(\tau; \triangle t), \tag{3.65}$$

where $A_c(\tau; \triangle t)$ gives the average output power associated with the channel as a function of the multipath delay $\tau = \tau_1 = \tau_2$ and the difference $\triangle t$ in observation time. The autocorrelation function (3.65) is based on the approximation that $\tau_1 = \tau_2$ when τ_1 and τ_2 satisfy $|\tau_1 - \tau_2| > B_u^{-1}$, since otherwise the receiver can't resolve the multipath components with these delays. In this case the two components are modeled as a single combined multipath component with delay $\tau \approx \tau_1 \approx \tau_2$ and autocorrelation $A_c(\tau; \triangle t)$.

[8] It is easily shown that the autocorrelation of the bandpass channel response $h(\tau, t)$ is $\mathbf{E}[h(\tau_1, t)h(\tau_2, t+\triangle t)] = .5\,\mathrm{Re}\{A_c(\tau_1, \tau_2; t, t+\triangle t)e^{j2\pi f_c \triangle t}\} + .5\,\mathrm{Re}\{\hat{A}_c(\tau_1, \tau_2; t, t+\triangle t)e^{j2\pi f_c (2t+\triangle t)}\}$, where $\hat{A}_c(\tau_1, \tau_2; t,\ t+\triangle t) = \mathbf{E}[c(\tau_1; t)c(\tau_2; t+\triangle t)]$. However, if $c(\tau, t)$ is WSS then $\hat{A}_c(\tau_1, \tau_2; t, t+\triangle t) = 0$, so $\mathbf{E}[h(\tau_1, t)h(\tau_2, t+\triangle t)] = .5\,\mathrm{Re}\{A_c(\tau_1, \tau_2; t, t+\triangle t)e^{j2\pi f_c \triangle t}\}$.

The *scattering function* for random channels is defined as the Fourier transform of $A_c(\tau; \triangle t)$ with respect to the $\triangle t$ parameter:

$$S_c(\tau,\ \rho) = \int_{-\infty}^{\infty} A_c(\tau,\ \triangle t) e^{-j2\pi\rho\triangle t} d\triangle t. \tag{3.66}$$

The scattering function characterizes the average output power associated with the channel as a function of the multipath delay τ and Doppler ρ. By contrast, for the deterministic channel, its scattering function $C(\tau,\ \rho)$ characterizes the complex channel gain as a function of multipath delay τ and Doppler ρ. An example scattering function for an indoor channel with the power density and delay spread normalized by their mean values is shown in Figure 3.12.

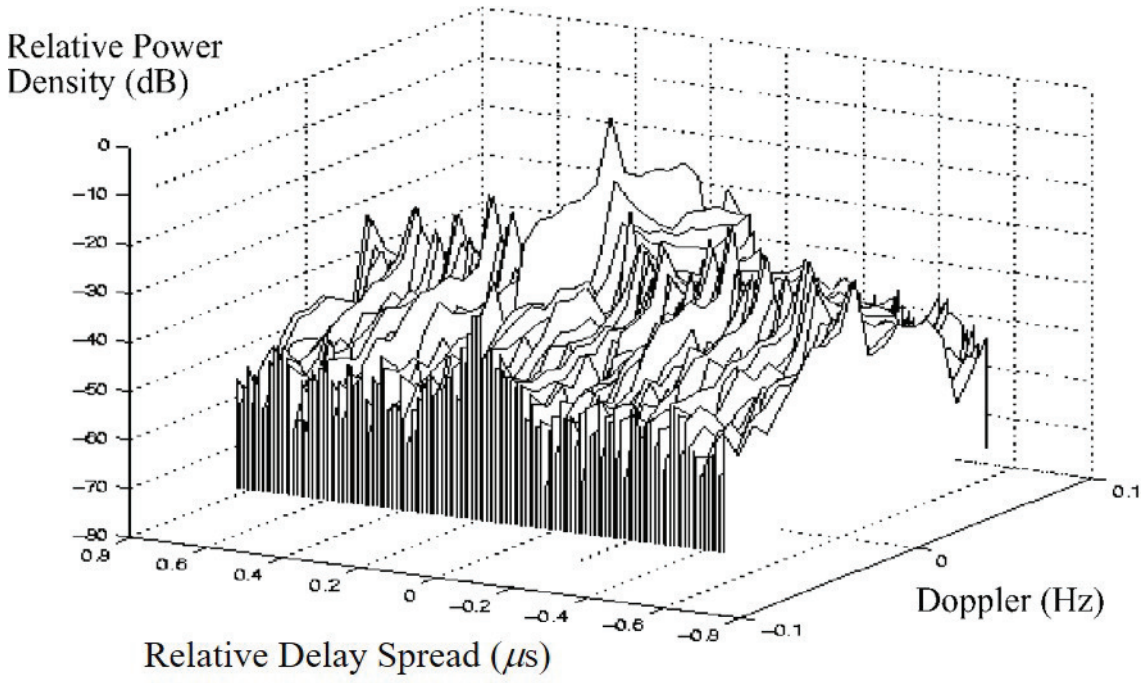


Figure 3.12: Scattering function with power spectral density and delay spread normalized by their mean values.

The most important characteristics of the wideband channel – including its power delay profile, coherence bandwidth, Doppler power spectrum, and coherence time–are derived from the channel autocorrelation $A_c(\tau,\ \triangle t)$ or the scattering function $S_c(\tau,\ \rho)$. These characteristics are described in subsequent sections.

3.3.2 Power Delay Profile

The *power delay profile* $A_c(\tau)$, also called the *multipath intensity profile*, is defined as the autocorrelation (3.65) with $\triangle t = 0 : A_c(\tau) \triangleq A_c(\tau, 0)$. The power delay profile represents the average power associated with a given multipath delay, and it is easily measured empirically. The average and rms delay spread are typically defined in terms of the power delay profile $A_c(\tau)$ as

$$\mu_{T_m} = \frac{\int_0^\infty \tau A_c(\tau) d\tau}{\int_0^\infty A_c(\tau) d\tau}, \tag{3.67}$$

and

$$\sigma_{T_m} = \sqrt{\frac{\int_0^\infty (\tau - \mu_{T_m})^2 A_c(\tau) d\tau}{\int_0^\infty A_c(\tau) d\tau}}. \tag{3.68}$$

Note that $A_c(\tau) \geq 0$ for all τ, so if we define the distribution p_{T_m} of the random delay spread T_m as

$$p_{T_m}(\tau) = \frac{A_c(\tau)}{\int_0^\infty A_c(\tau) d\tau} \tag{3.69}$$

then μ_{T_m} and σ_{T_m} are (respectively) the mean and rms values of T_m, relative to this distribution. Defining the distribution of T_m by (3.69) – or, equivalently, defining the mean and rms delay spread by (3.67) and (3.68), respectively – weighs the delay associated with a given multipath component by its relative power, so that weak multipath components contribute less to delay spread than strong ones. In particular, multipath components below the noise floor will not significantly affect these delay spread characterizations.

The time delay T where $A_c(\tau) \approx 0$ for $\tau \geq T$ can be used to roughly characterize the delay spread of the channel, and this value is often taken to be a small integer multiple of the rms delay spread. For example, we assume $A_c(\tau) \approx 0$ for $\tau > 3\sigma_{T_m}$. With this approximation, a linearly modulated signal with symbol period T_s experiences significant ISI if $T_s \ll \sigma_{T_m}$. Conversely, when $T_s \gg \sigma_{T_m}$ the system experiences negligible ISI. For calculations one can assume that $T_s \ll \sigma_{T_m}$ implies $T_s < \sigma_{T_m}/10$ and that $T_s \gg \sigma_{T_m}$ implies $T_s > 10\sigma_{T_m}$. If T_s is within an order of magnitude of σ_{T_m} then there will be some ISI, which may or may not significantly degrade performance, depending on the specifics of the system and channel. In later chapters we will study the performance degradation due to ISI in linearly modulated systems as well as ISI mitigation methods.

Although $\mu_{T_m} \approx \sigma_{T_m}$ in many channels with a large number of scatterers, the exact relationship between μ_{T_m} and σ_{T_m} depends on the shape of $A_c(\tau)$. For a channel with no LOS component and a small number of multipath components with approximately the same large delay, $\mu_{T_m} \gg \sigma_{T_m}$. In this case the large value of μ_{T_m} is a misleading metric of delay spread, since in fact all copies of the transmitted signal arrive at roughly the same time and the demodulator would synchronize to this common delay. It is typically assumed that the synchronizer locks to the multipath component at approximately the mean delay, in which case rms delay spread characterizes the time spreading of the channel.

Example 3.4: The power delay profile is often modeled as having a one-sided exponential distribution:

$$A_c(\tau) = \frac{1}{\overline{T}_m} e^{-\tau/\overline{T}_m}, \ \tau \geq 0. \tag{3.70}$$

Show that the average delay spread (3.67) is $\mu_{T_m} = \overline{T}_m$ and find the rms delay spread (3.68).

Solution: It is easily shown that $A_c(\tau)$ integrates to unity. The average delay spread is thus given by

$$\mu_{T_m} = \frac{1}{\overline{T}_m} \int_0^\infty \tau e^{-\tau/\overline{T}_m} d\tau = \overline{T}_m, \tag{3.71}$$

and the rms delay spread is

$$\sigma_{T_m} = \sqrt{\frac{1}{\overline{T}_m} \int_0^\infty \tau^2 e^{-\tau/\overline{T}_m} d\tau - \mu_{T_m}^2} = 2\overline{T}_m - \overline{T}_m = \overline{T}_m. \tag{3.72}$$

Thus, the average and rms delay spread are the same for exponentially distributed power delay profiles.

Example 3.5: Consider a wideband channel with multipath intensity profile

$$A_c(\tau) = \begin{cases} e^{-\tau/.00001} & 0 \leq \tau \leq 20\ \mu\text{s}, \\ 0 & \text{else.} \end{cases} \tag{3.73}$$

Find the mean and rms delay spreads of the channel and find the maximum symbol rate such that a linearly modulated signal transmitted through this channel does not experience ISI.

Solution: The average delay spread is

$$\mu_{T_m} = \frac{\int_0^{20\cdot 10^{-6}} \tau e^{-\tau/.00001} d\tau}{\int_0^{20\cdot 10^{-6}} e^{-\tau/.00001} d\tau} = 6.87\ \mu\text{s}. \tag{3.74}$$

The rms delay spread is

$$\sigma_{T_m} = \sqrt{\frac{\int_0^{20\cdot 10^{-6}} (\tau - \mu_{T_m})^2 e^{-\tau/.00001} d\tau}{\int_0^{20\cdot 10^{-6}} e^{-\tau/.00001} d\tau}} = 5.25\ \mu\text{s}. \tag{3.75}$$

We see in this example that the mean delay spread is roughly equal to its rms value. To avoid ISI we require linear modulation to have a symbol period T_s that is large relative to σ_{T_m}. Taking this to mean that $T_s > 10\sigma_{T_m}$ yields a symbol period of $T_s = 52.5\ \mu$s or a symbol rate of $R_s = 1/T_s = 19.04$ kilosymbols per second. This is a highly constrained symbol rate for many wireless systems. Specifically, for binary modulations where the symbol rate equals the data rate (bits per second, or bps), voice requires on the order of 32 kbps and data has orders of magnitude higher rates, with some systems operating at Gbps data rates.

3.3.3 Coherence Bandwidth

We can also characterize the time-varying multipath channel in the frequency domain by taking the Fourier transform of $c(\tau,\ t)$ with respect to τ. Specifically, define the random process

$$C(f;t) = \int_{-\infty}^{\infty} c(\tau;t) e^{-j2\pi f\tau} d\tau. \tag{3.76}$$

Because $c(\tau;t)$ is a complex zero-mean Gaussian random variable in t, the Fourier transform in (3.76) represents the sum[9] of complex zero-mean Gaussian random processes; hence $C(f;t)$ is also a zero-mean Gaussian random process that is completely characterized by its autocorrelation. Since $c(\tau;t)$ is WSS, its integral $C(f;t)$ is also. Thus, the autocorrelation of (3.76) is given by

$$A_C(f_1,\ f_2;\triangle t) = \mathbf{E}[C^*(f_1;t)C(f_2;t+\triangle t)]. \tag{3.77}$$

We can simplify $A_C(f_1,\ f_2;\triangle t)$ as follows:

[9] We can express the integral as a limit of a discrete sum.

$$
\begin{aligned}
A_C(f_1,\ f_2;\triangle t) &= \mathbf{E}\left[\int_{-\infty}^{\infty} c^*(\tau_1;t)e^{j2\pi f_1\tau_1}d\tau_1 \int_{-\infty}^{\infty} c(\tau_2;t+\triangle t)e^{-j2\pi f_2\tau_2}d\tau_2\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{E}[c^*(\tau_1;t)c(\tau_2;t+\triangle t)]e^{j2\pi f_1\tau_1}e^{-j2\pi f_2\tau_2}d\tau_1 d\tau_2 \\
&= \int_{-\infty}^{\infty} A_c(\tau,\ \triangle t)e^{-j2\pi(f_2-f_1)\tau}d\tau \\
&= A_C(\triangle f;\triangle t), \qquad (3.78)
\end{aligned}
$$

where $\triangle f = f_2 - f_1$ and the third equality follows from the WSS and US properties of $c(\tau;t)$. Thus, the autocorrelation of $C(f;t)$ in frequency depends only on the frequency difference $\triangle f$. The function $A_C(\triangle f;\triangle t)$ can be measured in practice by transmitting a pair of sinusoids through the channel that are separated in frequency by $\triangle f$ and then calculating their cross-correlation at the receiver for the time separation $\triangle t$.

If we define $A_C(\triangle f) \triangleq A_C(\triangle f;0)$ then, by (3.78),

$$
A_C(\triangle f) = \int_{-\infty}^{\infty} A_c(\tau)e^{-j2\pi\triangle f\tau}d\tau. \qquad (3.79)
$$

Thus $A_C(\triangle f)$ is the Fourier transform of the power delay profile. Because $A_C(\triangle f) = \mathbf{E}[C^*(f;t)C(f+\triangle f;t)]$ is an autocorrelation, it follows that the channel response is approximately independent at frequency separations $\triangle f$ where $A_C(\triangle f) \approx 0$. The frequency B_c where $A_C(\triangle f) \approx 0$ for all $\triangle f > B_c$ is called the *coherence bandwidth* of the channel. By the Fourier transform relationship between $A_c(\tau)$ and $A_C(\triangle f)$, if $A_c(\tau) \approx 0$ for $\tau > T$ then $A_C(\triangle f) \approx 0$ for $\triangle f > 1/T$. Hence, the minimum frequency separation B_c for which the channel response is roughly independent is $B_c \approx 1/T$, where T is typically taken to be the rms delay spread σ_{T_m} of $A_c(\tau)$. A more general approximation is $B_c \approx k/\sigma_{T_m}$, where k depends on the shape of $A_c(\tau)$ and the precise specification of coherence bandwidth. For example, Lee [24] has shown that $B_c \approx .02/\sigma_{T_m}$ approximates the range of frequencies over which channel correlation exceeds 0.9 whereas $B_c \approx .2/\sigma_{T_m}$ approximates the range of frequencies over which this correlation exceeds 0.5.

In general, if we are transmitting a narrowband signal with bandwidth $B \ll B_c$, then fading across the entire signal bandwidth is highly correlated; that is, the fading is roughly equal across the entire signal bandwidth. This is usually referred to as *flat fading*. On the other hand, if the signal bandwidth $B \gg B_c$, then the channel amplitude values at frequencies separated by more than the coherence bandwidth are roughly independent. Thus, the channel amplitude varies widely across the signal bandwidth. In this case the fading is called *frequency selective*. If $B \approx B_c$ then channel behavior is somewhere between flat and frequency-selective fading. Note that in linear modulation the signal bandwidth B is inversely proportional to the symbol time T_s, so flat fading corresponds to $T_s \approx 1/B \gg 1/B_c \approx \sigma_{T_m}$ – that is, the case where the channel experiences negligible ISI. Frequency-selective fading corresponds to $T_s \approx 1/B \ll 1/B_c = \sigma_{T_m}$, the case where the linearly modulated signal experiences significant ISI. Wideband signaling formats that reduce ISI, such as multicarrier modulation and spread spectrum, still experience frequency-selective fading across their entire signal bandwidth; this degrades performance, as will be discussed in Chapters 12 and 13.

We illustrate the power delay profile $A_c(\tau)$ and its Fourier transform $A_C(\triangle f)$ in Figure 3.13. This figure also shows two signal spectra superimposed on $A_C(\triangle f)$: a narrowband signal with bandwidth much less than B_c, and a wideband signal with bandwidth much greater than B_c. We see that the autocorrelation $A_C(\triangle f)$ is flat across the bandwidth of the narrowband signal, so this signal will experience flat fading or (equivalently) negligible ISI. The autocorrelation $A_C(\triangle f)$ goes to zero within the bandwidth of the wideband signal, which means that fading will be independent across different parts of the signal bandwidth; hence fading is frequency selective, and a linearly modulated signal transmitted through this channel will experience significant ISI.

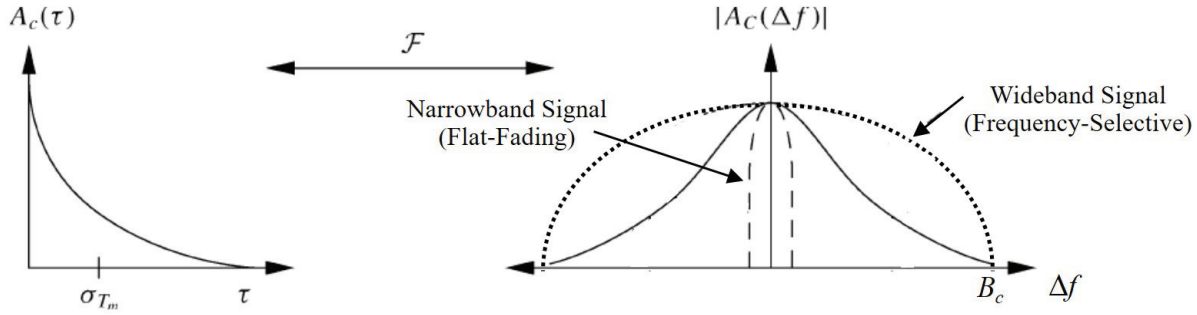


Figure 3.13: Power delay profile, rms delay spread, and coherence bandwidth.

Example 3.6: Consider an indoor channel with $\sigma_{T_m} = 30$ ns and a large macrocell channel with $\sigma_{T_m} = 300$ ns. Find the maximum symbol rate $R_s = 1/T_s$ for these environments such that a linearly modulated signal transmitted through them experiences negligible ISI.

Solution: We assume that negligible ISI requires $T_s \gg \sigma_{T_m}$ (i.e., $T_s \geq 10\sigma_{T_m}$). This translates into a symbol rate of $R_s = 1/T_s \leq .1/\sigma_{T_m}$. For $\sigma_{T_m} = 30$ ns this yields $R_s \leq 3.33$ Mbps and for $\sigma_{T_m} = 300$ ns this yields $R_s \leq 333$ Kbps. Note that high-performance wireless systems today can support Gbps data rates indoors and out. To maintain these data rates for a linearly modulated signal without severe performance degradation by ISI, some form of ISI mitigation is needed. Moreover, ISI is less severe in indoor than in outdoor systems owing to the former's lower delay spread values, which is why indoor systems tend to have higher data rates than outdoor systems.

3.3.4 Doppler Power Spectrum and Channel Coherence Time

The time variations of the channel that arise from transmitter or receiver motion cause a Doppler shift in the received signal. This Doppler effect can be characterized by taking the Fourier transform of $A_C(\triangle f; \triangle t)$ relative to $\triangle t$:

$$S_C(\triangle f; \rho) = \int_{-\infty}^{\infty} A_C(\triangle f; \triangle t) e^{-j2\pi\rho\triangle t} d\triangle t. \tag{3.80}$$

In order to characterize Doppler at a single frequency, we set $\triangle f$ to zero and then define $S_C(\rho) \stackrel{\triangle}{=} S_C(0; \rho)$. It is easily seen that

$$S_C(\rho) = \int_{-\infty}^{\infty} A_C(\triangle t) e^{-j2\pi\rho\triangle t} d\triangle t, \tag{3.81}$$

where $A_C(\triangle t) \triangleq A_C(\triangle f = 0; \triangle t)$. Note that $A_C(\triangle t)$ is an autocorrelation function defining how the channel impulse response decorrelates over time. In particular, $A_C(\triangle t = T) = 0$ indicates that observations of the channel impulse response at times separated by T are uncorrelated and therefore independent, since the channel impulse response is a Gaussian random process. We define the *channel coherence time* T_c to be the range of $\triangle t$ values over which $A_C(\triangle t)$ is approximately nonzero. Thus, the time-varying channel decorrelates after approximately T_c seconds. The function $S_C(\rho)$ is called the *Doppler power spectrum* of the channel. Since $S_C(\rho)$ is the Fourier

transform of an autocorrelation, it gives the PSD of the received signal as a function of Doppler ρ. The maximum ρ-value for which $|S_C(\rho)|$ is greater than zero is called the *Doppler spread* of the channel, denoted by B_D. By the Fourier transform relationship between $A_C(\triangle t)$ and $S_C(\rho)$, we have $B_D \approx 1/T_c$. If the transmitter and reflectors are all stationary and the receiver is moving with velocity v, then $B_D \leq v/\lambda = f_D$. Recall that in the narrowband fading model samples became independent at time $\triangle t = .4/f_D$, so in general $B_D \approx k/T_c$, where k depends on the shape of $S_c(\rho)$. We illustrate the Doppler power spectrum $S_C(\rho)$ and its inverse Fourier transform $A_C(\triangle_t)$ in Figure 3.14.

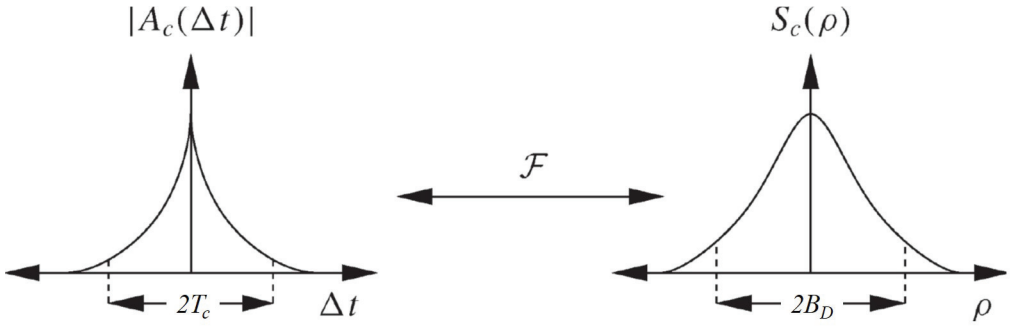


Figure 3.14: Doppler power spectrum, Doppler spread, and coherence time.

Example 3.7: For a channel with Doppler spread $B_D = 80$ Hz, find the time separation required in samples of the received signal in order for the samples to be approximately independent.

Solution: The coherence time of the channel is $T_c \approx 1/B_D = 1/80$, so samples spaced 12.5 ms apart are approximately uncorrelated. Thus, given the Gaussian properties of the underlying random process, these samples are approximately independent.

3.3.5 Transforms for Autocorrelation and Scattering Functions

From (3.80) we see that the scattering function $S_c(\tau;\rho)$ defined in (3.66) is the inverse Fourier transform of $S_C(\triangle f;\rho)$ in the $\triangle f$ variable. Furthermore $S_c(\tau;\rho)$ and $A_C(\triangle f;\triangle t)$ are related by the double Fourier transform

$$S_c(\tau;\rho) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_C(\triangle f;\triangle t)e^{-j2\pi\rho\triangle t}e^{j2\pi\tau\triangle f}d\triangle t\, d\triangle f. \tag{3.82}$$

The relationships among the four functions $A_C(\triangle f;\triangle t)$, $A_c(\tau;\triangle t)$, $S_C(\triangle f;\rho)$, and $S_c(\tau;\rho)$ are shown in Figure 3.15.

Empirical measurements of the scattering function for a given channel are often used to approximate the channel's delay spread, coherence bandwidth, Doppler spread, and coherence time. The delay spread for a channel with empirical scattering function $S_c(\tau;\rho)$ is obtained by computing the empirical power delay profile $A_c(\tau)$ from $A_c(\tau,\ \triangle t) = \mathcal{F}_\rho^{-1}[S_c(\tau;\rho)]$ with $\triangle t = 0$ and then computing the mean and rms delay spread from this power delay profile. The coherence bandwidth can then be approximated as $B_c \approx 1/\sigma_{T_m}$. Similarly, the Doppler spread B_D is approximated as the range of ρ values over which $S(0;\rho)$ is roughly nonzero, with the coherence time $T_c \approx 1/B_D$.

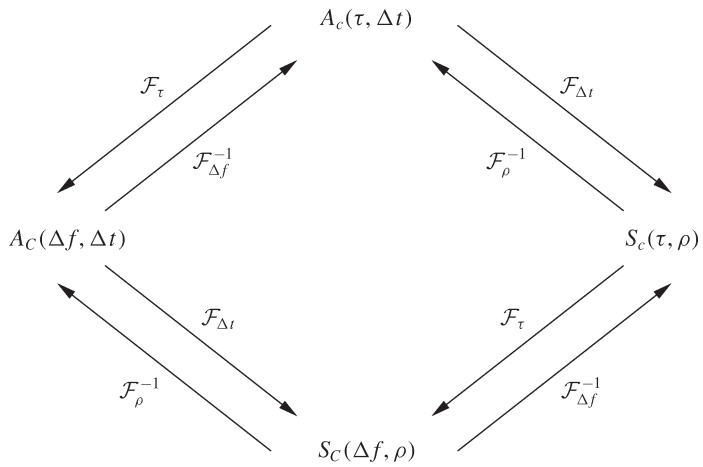


Figure 3.15: Fourier transform relationships.

To summarize the channel characteristics developed in this section, a signal transmitted over a wideband fading channel experiences time variations due to Doppler and frequency variations due to delay spread. If the transmitted signal is sufficiently wide in bandwidth (much greater than the coherence bandwidth) and long in duration (much longer than the coherence time) then it will experience the *full diversity* of the channel, meaning that the signal spans the channel variations over both time and frequency. Multicarrier and spread spectrum modulation techniques, described in Chapters 12 and 13, respectively, are designed to exploit this full diversity in order to mitigate the detrimental performance effects caused by time-varying fading and intersymbol interference.

3.4 Discrete-Time Model

Often the time-varying impulse response channel model is too complex for simple analysis. In this case a discrete-time approximation for the wideband multipath model can be used. This discrete-time model, developed by Turin in [25], is especially useful in the study of spread-spectrum systems and RAKE receivers (covered in Chapter 13). This discrete-time model is based on a physical propagation environment consisting of a composition of isolated point scatterers, as shown in Figure 3.16. In this model, the multipath components are assumed to form subpath clusters: incoming paths on a given subpath with approximate delay τ_i are combined, and incoming paths on different subpath clusters with delays τ_i and τ_j, where $|\tau_i - \tau_j| > 1/B_u$, can be resolved. Here B_u denotes the baseband signal bandwidth.

The channel model of (3.6) is modified to include a fixed number N of these subpath clusters as

$$c(\tau; t) = \sum_{i=0}^{N-1} \alpha_i(t) e^{-j\phi_i(t)} \delta(\tau - \tau_i(t)). \tag{3.83}$$

The statistics of the received signal for a given t are thus given by the statistics of $\{\tau_i\}_0^N$, $\{\alpha_i\}_0^N$, and $\{\phi_i\}_0^N$. Note that $c(\tau; t)$ in (3.83) has infinite bandwidth since it consists of a sum of delta functions. Since the model has infinite bandwidth, its discrete-time approximation associated with any finite sampling rate entails sampling distortion, as the channel's Nyquist rate is infinite.

With this sampling distortion in mind, the model of (3.83) can be approximated in discrete-time as follows. For a fixed t, the time axis is divided into M equal intervals of duration T such that $MT \geq \sigma_{T_m}$, where σ_{T_m} is the rms delay spread of the channel being approximated. The subpaths are restricted to lie in one of the M

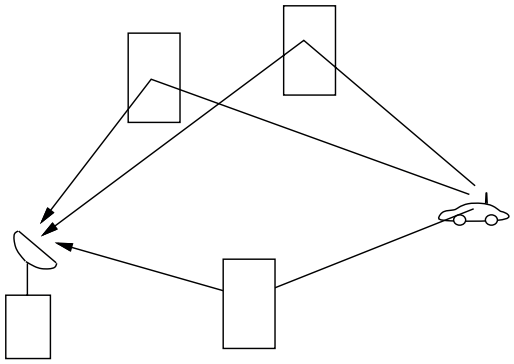


Figure 3.16: Point scatterer channel model.

time-interval bins, as shown in Figure 3.17. The multipath spread of this discrete model is MT, and the resolution between paths is T. This resolution is based on the transmitted signal bandwidth: $T \approx 1/B_u$. The statistics for the ith bin are r_i, $1 \leq i \leq M$, a binary indicator of the existence of a multipath component in the ith bin: so $r_i = 1$ if there is a multipath component in the ith bin and 0 otherwise. If $r_i = 1$ then $(a_i,\ \theta_i)$, the amplitude and phase corresponding to this multipath component, follow an empirically determined distribution. This distribution is obtained by sample averages of $(a_i,\ \theta_i)$ for each i at different locations in the propagation environment. The empirical distribution of $(a_i,\ \theta_i)$ and $(a_j,\ \theta_j), i \neq j$, is generally different; it may correspond to the same family of fading but with different parameters (e.g., Rician fading with different K factors) or to different fading distributions altogether (e.g., Rayleigh fading for the ith bin, Nakagami fading for the jth bin).

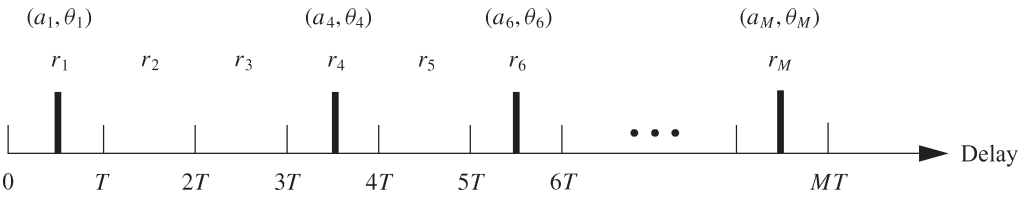


Figure 3.17: Discrete-time approximation.

This completes our statistical model of the discrete-time approximation for a single snapshot. A sequence of profiles will model the signal over time as the channel impulse response changes – for example, the impulse response seen by a receiver moving at some nonzero velocity through a city. Thus, the model must include not only the first-order statistics of $(\tau_i, \alpha_i, \phi_i)$ for each profile (equivalently, each t) but also the temporal and spatial correlations (assumed to be Markov) between them. More details on the model and the empirically derived distributions for N and for $(\tau_i, \alpha_i, \phi_i)$ can be found in [25].

3.5 MIMO Channel Models

Multiple antennas at the transmitter and/or receiver are common in wireless systems because of their diversity and capacity benefits. These MIMO systems with multiple antennas require channel models that characterize both spatial and temporal characteristics of the channel. For a channel with M transmit antennas and N receive

antennas, the MIMO channel can be written as a matrix

$$H(t,\tau) = \begin{pmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & \cdots & h_{1M}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & \cdots & h_{2M}(t,\tau) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1}(t,\tau) & h_{N2}(t,\tau) & \cdots & h_{NM}(t,\tau) \end{pmatrix} \tag{3.84}$$

where $h_{ij}(t,\tau)$ is the time-varying impulse response between the jth transmit antenna and the ith receiver antenna as given by (3.6). When the channel is narrowband and time-invariant, each of these terms reduces to a complex channel gain h_{ij}. In this case the MIMO channel model consists of the distributions and cross correlations of these random channel gains. The class of MIMO channel models that capture only these gains is often referred to as the *analytical* MIMO channel model. The complex channel gains h_{ij} in MIMO systems are often assumed to be i.i.d. zero mean, unit variance, circularly symmetric Gaussian random variables, meaning that a fixed phase rotation has no impact on the distribution. This distribution for the MIMO channel gains is based on a physical environment where there is a sufficient number of scatterers to yield independent fading over each transmit-receive signal path. Environments with this characteristic are said to exhibit *rich scattering*. Models for more general environments with antenna gains that are correlated or dependent include the Rice, Kronecker, UIU, Weichselberger, and keyhole models [26]. For each transmit-receive antenna pair in the MIMO system, the single-antenna wideband fading model described in Section 3.3 with channel gain h_{ij} can be used to capture the effects of time variations and delay spread. However, this modeling approach does not capture dependencies across all three MIMO dimensions of space, time, and frequency.

To capture dependencies across all three of these dimensions, we extend the time-varying channel impulse response model (3.6) to take into account multiple transmit and receive antennas. As in the single-antenna case, the propagation environment assumes multiple scatterers that generate the multipath components [26, 27]. The locations of the scatterers at time t dictate the AoA $\theta_i(t)$ at the receiver array and the angle-of-departure (AoD) $\psi_i(t)$ at the transmitter array for the ith multipath component. Models can be either two-dimensional or three-dimensional. Since knowledge of the environment via the angles of arrival and departure is used in this model, it is sometimes referred to as the *parametric* MIMO channel model to contrast it with the analytical model.

The equivalent lowpass time-varying impulse response model (3.6) is extended to incorporate these angles of arrival and departure as follows:

$$c(\tau,t,\theta,\psi) = \sum_{i=0}^{N(t)-1} \alpha_i(t) e^{-j\phi_i(t)} \delta(\tau - \tau_i(t)) \delta(\theta - \theta_i) \delta(\psi - \psi_i) \tag{3.85}$$

where $N(t)$, $\alpha_i(t)$, $\phi_i(t)$ and $\tau_i(t)$ are defined as in (3.6). This time-varying impulse response is zero for all values of the AoA θ that do not correspond to one of the incoming multipath components at time t associated with the physical environment, and similarly for all values of the AoD ψ. In three dimensions the angles of arrival and departure of each multipath component depends on the elevation and azimuth angles, the polarization of antenna elements, as well as the scatterer locations. Survey papers on such 3-D channel models include [28, 29].

Chapter 3 Problems

1. Consider a two-ray channel consisting of a direct ray plus a ground-reflected ray where the transmitter is a fixed base station at height h and the receiver is mounted on a truck (also at height h). The truck starts next to the base station and moves away at velocity v. Assume signal attenuation on each path follows a free-space path loss model. Find the time-varying channel impulse at the receiver for transmitter-receiver separation $d = vt$ sufficiently large for the length of the reflected path to be approximated by $d_{11} + d_{12} \approx d + 2h^2/d$. Assume a carrier frequency $f = 2$ GHz, unit gain antennas at the transmitter and receiver, and reflection coefficient $R = -1$.

2. Find a formula for the multipath delay spread T_m for a two-ray channel model. Find a simplified formula when the transmitter-receiver separation is relatively large. Compute T_m for $h_t = 10$ m, $h_r = 4$ m, and $d = 100$ m.

3. Consider a time-invariant indoor wireless channel with LOS component at delay 23 nsec, a multipath component at delay 48 ns, and another multipath component at delay 67 ns. Find the delay spread assuming the demodulator synchronizes to the LOS component. Repeat assuming that the demodulator synchronizes to the first multipath component.

4. Show that the minimum value of $f_c\tau_i$ for a system at $f_c = 1$ GHz with a fixed transmitter and a receiver separated by more than 10 m from the transmitter is much greater than 1.

5. Prove that for X and Y independent zero-mean Gaussian random variables with variance σ^2, $Z = \sqrt{X^2 + Y^2}$ is Rayleigh-distributed and Z^2 is exponentially-distributed.

6. Assume a Rayleigh fading channel with average signal power $2\sigma^2 = -80$ dBm. What is the power outage probability of this channel relative to the threshold $P_o = -95$ dBm? How about $P_o = -90$ dBm?

7. Suppose we have an application that requires a power outage probability of .01 for the threshold $P_o = -80$ dBm, For Rayleigh fading, what value of the average signal power is required?

8. Assume a Rician fading channel with $2\sigma^2 = -80$ dBm and a target power of $P_o = -80$ dBm. Find the outage probability assuming that the LOS component has average power $s^2 = -80$ dBm.

9. This problem illustrates that the tails of the Rician distribution can be quite different than its Nakagami approximation. Plot the cumulative distribution function (CDF) of the Rician distribution for $K = 1, 5, 10$ and the corresponding Nakagami distribution with $m = (K + 1)^2/(2K + 1)$. In general, does the Rician distribution or its Nakagami approximation have a larger outage probability $p(\gamma < x)$ for x large?

10. The TWDP narrowband fading model, described in Section 3.2.2, has two dominant multipath components as well as a large number of lower-power multipath components. All multipath components are assumed to have i.i.d. uniformly-distributed phase. Assuming the non-dominant multipath components approximate the Rayleigh fading model, the received signal V_r is given by

$$V_r = V_1 e^{j\phi_1} + V_2 e^{j\phi_2} + X + jY,$$

where V_1 and V_2 are non-negative constants, ϕ_i and ϕ_2 are i.i.d. and uniformly distributed over $[0, 2\pi]$, and X and Y are i.i.d. Gauss-distributed random variables with mean zero and variance σ^2. The TWDP model

thus reduces to the Rician fading model if $V_1 > 0, V_2 = 0$ and to the Rayleigh fading model if $V_1 = V_2 = 0$. The model is parameterized in terms of K and Δ where,

$$\begin{aligned} K &= \frac{V_1^2 + V_2^2}{2\sigma^2} \\ \Delta &= \frac{2V_1V_2}{V_1^2 + V_2^2} \end{aligned}$$

(a) Find the range of parameters K and Δ.

(b) Describe the channel model that corresponds to the following TWDP parameters:

- $K \to \infty, \Delta \to 0$
- $K \to \infty, \Delta \to 1$
- $K \to 0$
- $K > 0, \Delta \to 0$

(c) Show that the received signal given that the phase difference between the 2 dominant components is constant, i.e., $\phi_1 - \phi_2 = \psi$ for a constant ψ, consists of one dominant component and the multiple diffuse lower-power multipath components. This implies that the TWDP model reduces to the Rician fading model when the phase difference between the LOS components in constant. Find the parameter $\bar{K}(\psi, \Delta, K)$ of the equivalent Rician model as a function of ψ, Δ and K.

(d) Show that $\psi \pmod{2\pi} \sim \mathcal{U}[0, 2\pi]$. Use this to show that

$$f_{\text{GTR}}(r \mid K, \Delta) = \frac{1}{2\pi} \int_0^{2\pi} f_{\text{Rice}}(r \mid \bar{K}(\psi, \Delta, K)) \mathrm{d}\psi,$$

where $f_{\text{TWDP}}(r \mid \Sigma)$ is the pdf of the received signal amplitude of the TWDP channel fading model with parameters Σ and $f_{\text{Rice}}(r \mid \Sigma)$ is the pdf of the received signal amplitude of the TWDP channel fading model with parameters Σ.

11. In order to improve the performance of cellular systems, multiple base stations can receive the signal transmitted from a given mobile unit and combine these multiple signals either by selecting the strongest one or summing the signals together, perhaps with some optimized weights. This typically increases SNR and reduces the effects of shadowing. Combining of signals received from multiple base stations is called *macro-diversity*, and in this problem we explore the benefits of this technique. Diversity will be covered in more detail in Chapter 7.

Consider a mobile at the midpoint between two base stations in a cellular network. The received signals (in dBW) from the base stations are given by

$$P_{r,1} = W + Z_1,$$

$$P_{r,2} = W + Z_2,$$

where $Z_{1,2}$ are $\mathcal{N}(0, \sigma^2)$ random variables. We define outage with macrodiversity to be the event that both $P_{r,1}$ and $P_{r,2}$ fall below a threshold T. The term W is the received signal power while Z_1 and Z_2 are independent noise terms.

(a) If Z_1 and Z_2 are independent, show that the outage probability is given by

$$P_{out} = [Q(\Delta/\sigma)]^2,$$

where $\Delta = W - T$ is the fade margin at the mobile's location.

(b) Now suppose Z_1 and Z_2 are correlated in the following way:

$$Z_1 = a\,Y_1 + b\,Y,$$

$$Z_2 = a\,Y_2 + b\,Y,$$

where Y, Y_1, Y_2 are independent $\mathcal{N}(0, \sigma^2)$ random variables, and a, b are such that $a^2 + b^2 = 1$. Show that

$$P_{out} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \left[Q\left(\frac{\Delta + by\sigma}{|a|\sigma} \right) \right]^2 e^{-y^2/2} dy.$$

(c) Compare the outage probabilities of (b) and (c) for the special case of $a = b = 1/\sqrt{2}$, $\sigma = 8$ and $\Delta = 5$ (this will require a numerical integration).

12. The goal of this problem is to develop a Rayleigh fading simulator for a mobile communications channel using the method based on filtering Gaussian processes based on the in-phase and quadrature PSDs described in 3.2.1. In this problem you must do the following:

(a) Develop simulation code to generate a signal with Rayleigh fading amplitude over time. Your sample rate should be at least 1000 samples/sec, the average received envelope should be 1, and your simulation should be parameterized by the Doppler frequency f_D. Matlab is the easiest way to generate this simulation, but any simulation tool can be used.

(b) Write a description of your simulation that clearly explains how your code generates the fading envelope using a block diagram and any necessary equations.

(c) Provide plots of received amplitude (dB) vs. time for $f_D = 1, 10, 100$ Hz. over 2 seconds.

13. For a Rayleigh fading channel with average power $\overline{P}_r = 30$ dB and Doppler frequency $f_D = 10$ Hz, compute the average fade duration for target fade values $P_0 = 0$ dB, $P_0 = 15$ dB, and $P_0 = 30$ dB.

14. Derive a formula for the average length of time a Rayleigh fading process with average power $\overline{P}_r$ stays **above** a given target fade value P_0. Evaluate this average length of time for $\overline{P}_r = 20$ dB, $P_0 = 25$ dB, and $f_D = 50$ Hz and compare with the average fade depth below P_0 associated with the same parameters.

15. Assume a Rayleigh fading channel with average power $\overline{P}_r = 10$ dB and Doppler $f_D = 80$ Hz. We would like to approximate the channel using a finite state Markov model with eight states and time interval $T = 10$ ms. The regions R_j corresponds to $R_1 = \{\gamma : -\infty \leq \gamma \leq -10 \text{ dB}\}$, $R_2 = \{\gamma : -10\text{dB} \leq \gamma \leq 0 \text{ dB}\}$, $R_3 = \{\gamma : 0\text{dB} \leq \gamma \leq 5 \text{ dB}\}$, $R_4 = \{\gamma : 5\text{dB} \leq \gamma \leq 10 \text{ dB}\}$, $R_5 = \{\gamma : 10\text{dB} \leq \gamma \leq 15 \text{ dB}\}$, $R_6 = \{\gamma : 15\text{dB} \leq \gamma \leq 20 \text{ dB}\}$, $R_7 = \{\gamma : 20\text{dB} \leq \gamma \leq 30 \text{ dB}\}$, $R_8 = \{\gamma : 30\text{dB} \leq \gamma \leq \infty \}$. Find the transition probabilities between each region for this model.

16. Consider the following channel scattering function obtained by sending a 900 MHz sinusoidal input into the channel:

$$S(\tau, \rho) = \begin{cases} \alpha_1 \delta(\tau) & \rho = 70\text{Hz}. \\ \alpha_2 \delta(\tau - .022\mu\text{s}) & \rho = 49.5\text{Hz}. \\ 0 & \text{else} \end{cases}$$

where α_1 and α_2 are determined by path loss, shadowing, and multipath fading. Clearly this scattering function corresponds to a 2-ray model. Assume the transmitter and receiver used to send and receive the sinusoid are located eight meters above the ground.

(a) Find the distance and velocity between the transmitter and receiver.

(b) For the distance computed in part (a), is the path loss as a function of distance proportional to d^{-2} or d^{-4}? *Hint: use the fact that the channel is based on a 2-ray model.*

(c) Does a 30 KHz voice signal transmitted over this channel experience flat or frequency-selective fading?

17. In this problem we examine the narrowband approximation in the two-ray model. Consider the two-ray model discussed in Chapter 2.4 with reflection coefficient $R = -1$, free space path loss for each ray and carrier frequency 2.4 GHz. Both the transmitter and the receiver are 40 m high.

(a) For a separation distance of 20 m, find the impulse response of the equivalent baseband channel in the time domain. What is the delay spread T_m?

(b) Repeat the computation of the impulse response and delay spread for a separation distance of 2000 m.

(c) Plot the baseband channel output in both the time and frequency domain in each of the above two channels for channel input $\text{sinc}(ft) \triangleq \frac{\sin(\pi ft)}{\pi ft}$ and $f = 20$ MHz. Is there ISI in either channel? Find the coherence bandwidth of the channel in each case and whether this bandwidth implies flat or frequency selective fading.

18. Consider a wideband channel characterized by the autocorrelation function

$$A_c(\tau, \Delta t) = \begin{cases} \text{sinc}(W\Delta t) & 0 \le \tau \le 10\mu\text{sec.} \\ 0 & \text{else} \end{cases},$$

where $W = 100$Hz and $\text{sinc}(x) = \sin(\pi x)/(\pi x)$.

(a) Does this channel correspond to an indoor channel or an outdoor channel, and why?

(b) Sketch the scattering function of this channel.

(c) Compute the channel's average delay spread, rms delay spread, and Doppler spread.

(d) Over approximately what range of data rates will a signal transmitted over this channel exhibit frequency-selective fading?

(e) Would you expect this channel to exhibit Rayleigh or Rician fading statistics, and why?

(f) Assuming that the channel exhibits Rayleigh fading, what is the average length of time that the signal power is continuously below its average value.

(g) Assume a system with narrowband binary modulation sent over this channel. Your system has error correction coding that can correct two simultaneous bit errors. Assume also that you always make an error if the received signal power is below its average value, and never make an error if this power is at or above its average value. If the channel is Rayleigh fading then what is the maximum data rate that can be sent over this channel with error-free transmission, making the approximation that the fade duration never exceeds twice its average value.

19. Let a scattering function $S(\tau, \rho)$ be nonzero over $0 \le \tau \le .1$ ms and $-.1 \le \rho \le .1$ Hz. Assume that the power of the scattering function is approximately uniform over the range where it is nonzero.

(a) What are the multipath spread and the Doppler spread of the channel?

(b) Suppose you input to this channel two sinusoids $\cos(2\pi f_1 t)$ and $\cos(2\pi f_2 t)$ separated in frequency by $\Delta f = |f_1 - f_2|$. What is the minimum value of Δf for which the channel response to the first sinusoid is approximately independent of the channel response to the second sinusoid.

(c) For two sinusoidal inputs to the channel $u_1(t) = \sin(2\pi ft)$ and $u_2(t) = \sin(2\pi f(t + \Delta t))$, find the minimum value of Δt for which the channel response to $u_1(t)$ is approximately independent of the channel response to $u_2(t)$.

(d) Will this channel exhibit flat fading or frequency-selective fading for a typical voice channel with a 3 KHz bandwidth? How about for a cellular channel with a 30 KHz bandwidth?

20. Consider a MIMO channel $\mathbf{H}$ with a time-invariant complex gain h_{ij} between the jth transmit antenna and the ith receive antenna. A MIMO channel is called a *keyhole* channel when the channel matrix $\mathbf{H}$ can be written as

$$\mathbf{H} = h_r h_t^H$$

where h_t is an $M \times 1$ vector of i.i.d. zero-mean unit variance complex Gaussian random variables, and h_r is an $N \times 1$ vector of i.i.d. zero-mean unit variance complex Gaussian random variables. Thus, h_{ij} is the product of two independent complex Gaussian random variables.

(a) Show that the power of h_{ij} has a distribution $p(x) = 2K_0(2\sqrt{x})$ where K_0 is the modified Bessel function of the second kind.

(b) Show that the matrix $\mathbf{H}$ has rank 1, i.e. show that all of its columns are linearly dependent, and the same for its rows.

The fact that the matrix $\mathbf{H}$ is rank 1 is what gives rise to the name keyhole channel, since the MIMO channel's multiple spatial dimensions reduce to a single dimension, as if the signal between the transmit and receive antenna arrays travels through a keyhole.

Bibliography

[1] T. S. Rappaport, *Wireless Communications – Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

[2] D. Parsons, *The Mobile Radio Propagation Channel*, Wiley, New York, 1994.

[3] R. S. Kennedy, *Fading Dispersive Communication Channels*, Wiley, New York, 1969.

[4] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell System Tech. J.*, pp. 957–1000, July/August 1968.

[5] W. C. Jakes, Jr., *Microwave Mobile Communications*, Wiley, New York, 1974.

[6] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*: *A Unified Approach to Performance Analysis*, 2nd Ed., Wiley, New York, 2004.

[7] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[8] J. Zhang, C. Pan, F. Pei, G. Liu, and X. Cheng, "Three-dimensional fading channel models: a survey of elevation angle research,"*IEEE Commun. Mag*., vol. 52, no. 6, pp. 218-226, 2014.

[9] M. Pätzold, *Mobile Fading Channels*, Wiley, New York, 2002.

[10] J. Cavers, *Mobile Channel Characteristics*, Springer 2000.

[11] S. O. Rice, "Mathematical analysis of random noise," *Bell System Tech. J*., pp. 282–333, July 1944, and pp. 46–156, January 1945.

[12] G. D. Durgin, T. S. Rappaport and D. A. de Wolf, "New analytical models and probability density functions for fading in wireless communications," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 1005-1015, June 2002.

[13] J. M. Romero-Jerez, F. J. Lopez-Martinez, J. F. Paris and A. J. Goldsmith, "The fluctuating two-ray fading model: Statistical characterization and performance analysis," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4420-4432, July 2017.

[14] F. Babich, G. Lombardi, and E. Valentinuzzi, "Variable order Markov modeling for LEO mobile satellite channels," *Elec. Lett*., pp. 621–3, April 1999.

[15] A. M. Chen and R. R. Rao, "On tractable wireless channel models," *Proc. Internat. Sympos. Pers*., *Indoor, Mobile Radio Commun*., pp. 825–30, September 1998.

[16] H. S. Wang and N. Moayeri, "Finite-state Markov channel – A useful model for radio communication channels," *IEEE Trans. Veh. Tech*., pp. 163–71, February 1995.

[17] C. C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun*., pp. 2032–40, December 2000.

[18] C. Pimentel and I. F. Blake, "Modeling burst channels using partitioned Fritchman's Markov models," *IEEE Trans. Veh. Tech*., pp. 885–99, August 1998.

[19] Y. L. Guan and L. F. Tumer, "Generalised FSMC model for radio channels with correlated fading," *IEE Proc. Commun*., pp. 133–7, April 1999.

[20] M. Chu and W. Stark, "Effect of mobile velocity on communications in fading channels," *IEEE Trans. Veh. Tech*., pp. 202–10, January 2000.

[21] C. Komninakis and R. D. Wesel, "Pilot-aided joint data and channel estimation in flat correlated fading," *Proc. IEEE Globecom Conf*., pp. 2534–9, November 1999.

[22] M. Peleg, S. Shamai (Shitz), and S. Galan, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channels," *IEE Proc. Commun*., pp. 87–95, April 2000.

[23] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst*., pp. 360–93, December 1963.

[24] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*, McGraw-Hill, New York, 1989.

[25] G. L. Turin, "Introduction to spread spectrum antimultipath techniques and their application to urban digital radio," *Proc. IEEE*, pp. 328–53, March 1980.

[26] R. W. Heath, Jr. and A. Lozano, *Foundations of MIMO Communication,* Cambridge University Press, 2019.

[27] R. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications Magazine*., pp. 10–22, February 1998.

[28] U. Karabulut, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Spatial and temporal channel characteristics of 5G 3D channel model with beamforming for user mobility investigations," *IEEE Communications Magazine*, Special Issue on Channel Models and Measurements for 5G, Dec. 2018.

[29] J. Zhang, C. Pan, F. Pei, G. Liu, and X. Cheng, "Three-dimensional fading channel models: a survey of elevation angle research," *IEEE Communications Magazine*, June 2014.

Chapter 4

Capacity of Wireless Channels

The growing demand for wireless communication makes it important to determine the capacity limits of the underlying channels for these systems. These capacity limits dictate the maximum data rates that can be transmitted over wireless channels with asymptotically small error probability, assuming no constraints on delay or complexity of the encoder and decoder. The mathematical theory of communication underlying channel capacity was pioneered by Claude Shannon in the late 1940s. This theory is based on the notion of mutual information between the input and output of a channel [1, 2, 3]. In particular, Shannon defined channel capacity as the channel's mutual information maximized over all possible input distributions. The significance of this mathematical construct was Shannon's coding theorem and its converse. The coding theorem proved that a code did exist that could achieve a data rate close to capacity with negligible probability of error. The converse proved that any data rate higher than capacity could not be achieved without an error probability bounded away from zero. Shannon's ideas were quite revolutionary at the time: the high data rates he predicted for telephone channels, and his notion that coding could reduce error probability without reducing data rate or causing bandwidth expansion. In time, sophisticated modulation and coding technology validated Shannon's theory and so, on telephone lines today, we achieve data rates very close to Shannon capacity with very low probability of error. These sophisticated modulation and coding strategies are treated in Chapters 5 and 8, respectively.

In this chapter we examine the capacity of a single-user wireless channel where transmitter and/or receiver have a single antenna. The capacity of single-user MIMO systems where the transmitter and receiver both have multiple antennas is treated in Chapter 10 and that of multiuser systems in Chapter 14. We will discuss capacity for channels that are both time invariant and time varying. We first look at the well-known formula for capacity of a time-invariant additive white Gaussian noise (AWGN) channel and then consider capacity of time-varying flat fading channels. Unlike the AWGN case, here the capacity of a flat fading channel is not given by a single formula because capacity depends on what is known about the time-varying channel at the transmitter and/or receiver. Moreover, for different channel information assumptions there are different definitions of channel capacity, depending on whether capacity characterizes the maximum rate averaged over all fading states or the maximum constant rate that can be maintained in all fading states (with or without some probability of outage).

We will first consider flat fading channel capacity where only the fading distribution is known at the transmitter and receiver. Capacity under this assumption is typically difficult to determine and is only known in a few special cases. Next we consider capacity when the channel fade level is known at the receiver only (via receiver estimation) or when the channel fade level is known at both the transmitter and the receiver (via receiver estimation and transmitter feedback). We will see that the fading channel capacity with channel fade level information at both the transmitter and receiver is achieved when the transmitter adapts its power, data rate, and coding scheme to the channel variation. The optimal power allocation in this case is a "water-filling" in time, where power and data rate are increased when channel conditions are favorable and decreased when channel conditions are not favorable.

We will also treat capacity of frequency-selective fading channels. For time-invariant frequency-selective channels the capacity is known and is achieved with an optimal power allocation that water-fills over frequency instead of time. The capacity of a time-varying frequency-selective fading channel is unknown in general. However, this channel can be approximated as a set of independent parallel flat fading channels whose capacity is the sum of capacities on each channel with power optimally allocated among the channels. The capacity of such a channel is known and the capacity-achieving power allocation water-fills over both time and frequency.

We will consider only discrete-time systems in this chapter. Most continuous-time systems can be converted to discrete-time systems via sampling, and then the same capacity results hold. However, care must be taken in choosing the appropriate sampling rate for this conversion, since time variations in the channel may increase the sampling rate required to preserve channel capacity [4].

4.1 Capacity in AWGN

Consider a discrete-time AWGN channel with channel input/out relationship $y[i] = x[i] + n[i]$, where $x[i]$ is the channel input at time i, $y[i]$ is the corresponding channel output, and $n[i]$ is a white Gaussian noise random process. Assume a channel bandwidth B and received signal power P. The received signal-to-noise ratio (SNR) – the power in $x[i]$ divided by the power in $n[i]$ – is constant and given by $\gamma = P/(N_0 B)$, where $N_0/2$ is the power spectral density (PSD) of the noise. The capacity of this channel is given by Shannon's well-known formula [1]:

$$C = B \log_2(1 + \gamma), \tag{4.1}$$

where the capacity units are bits per second (bps). Shannon's coding theorem proves that a code exists that achieves data rates arbitrarily close to capacity with arbitrarily small probability of bit error. The converse theorem shows that any code with rate $R > C$ has a probability of error bounded away from zero.

Shannon capacity for AWGN channels provides interesting insights along the asymptotic regimes of bandwidth and power. In particular, it is easily shown that

$$\lim_{B\to\infty} B \log_2 \left(1 + \frac{P}{N_0 B}\right) = \log_2(e)\frac{P}{N_0}. \tag{4.2}$$

This regime is called the *large bandwidth regime* for capacity. Interestingly, in this regime the capacity no longer depends on the bandwidth B. That is because there is not sufficient power to spread over the large amount of bandwidth available, so as the bandwidth grows large, there is diminishing return in capacity growth. The *low power regime* captures a similar scenario, where there is not sufficient power to utilize the available bandwidth. In particular, using the approximation that $\log(1 + x) \approx x$ for x small yields that, for P small,

$$B \log_2 \left(1 + \frac{P}{N_0 B}\right) \approx \log_2(e)\frac{P}{N_0}. \tag{4.3}$$

Here we also see that capacity no longer depends on the channel bandwidth. Another interesting regime for capacity is the *high power regime*, where P grows large. We now consider how capacity scales with bandwidth in this regime. Using the high-power approximation $1 + P/(N_0 B) \approx P/(N_0 B)$, we get the following ratio of AWGN channel capacity $C(kB)$ given bandwidth kB compared with $C(B)$ for any constant $k > 1$ in the limit of asymptotically large power:

$$\lim_{P\to\infty} \frac{C(kB)}{C(B)} = \lim_{P\to\infty} \frac{kB \log_2\left(1 + \frac{P}{N_0 kB}\right)}{B \log_2\left(1 + \frac{P}{N_0 B}\right)} \approx \lim_{P\to\infty} \frac{kB \log_2\left(\frac{P}{N_0 kB}\right)}{B \log_2\left(\frac{P}{N_0 B}\right)} = \lim_{P\to\infty} \frac{k\left[\log_2\left(\frac{P}{N_0 B}\right) + \log_2(1/k)\right]}{\log_2\left(\frac{P}{N_0 B}\right)} = k. \tag{4.4}$$

Thus, in the high power regime, scaling the bandwidth up by a factor of k leads to the same increase in capacity. That is because there is sufficient power to fully utilize the capacity of the additional bandwidth.

Shannon's coding theorem and its converse are proved using the concept of mutual information between the channel input and output. For a discrete memoryless time-invariant channel with random input x and random output y, the channel's *mutual information* is defined as

$$I(X;Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y) \, \log \left(\frac{p(x,y)}{p(x)p(y)} \right), \tag{4.5}$$

where the sum is taken over all possible input and output pairs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ for $\mathcal{X}$ and $\mathcal{Y}$ the discrete input and output alphabets. The log function is typically with respect to base 2, in which case the units of mutual information are bits per second. Mutual information can also be written in terms of the *entropy* in the channel output $y \,|\, x$ as $I(X;Y) = H(Y) - H(Y \,|\, X)$, where $H(Y) = -\sum_{y \in \mathcal{Y}} p(y) \log p(y)$ and $H(Y \,|\, X) = -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, \ y) \log p(y \,|\, x)$. Shannon proved that channel capacity equals the mutual information of the channel maximized over all possible input distributions:

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} \sum_{x,y} p(x,y) \, \log \left(\frac{p(x,y)}{p(x)p(y)} \right). \tag{4.6}$$

For the AWGN channel, the sum in (4.6) becomes an integral over continuous alphabets and the maximizing input distribution is Gaussian, which results in the channel capacity given by (4.1). For channels with memory, mutual information and channel capacity are defined relative to input and output sequences x^n and y^n. More details on channel capacity, mutual information, the coding theorem, and its converse can be found in [2, 5, 6].

The proofs of the coding theorem and its converse place no constraints on the complexity or delay of the communication system. Indeed, the capacity-achieving strategy requires a code with an asymptotically large blocklength to drive the error probability to zero. Such codes require asymptotically large transmission energy and asymptotically large delay and complexity to decode. Since practical decoders are constrained in delay and complexity, Shannon capacity is generally used as an upper bound on the data rates that can be achieved under real system constraints. The exact capacity penalty due to finite blocklength codes was derived in [7] as a function of blocklength, error probability, and the *channel dispersion*, which measures the stochastic variability of the channel relative to a deterministic channel with the same capacity.

At the time that Shannon developed his theory of information, data rates over standard telephone lines were on the order of 100 bps. Thus, it was believed that Shannon capacity, which predicted speeds of roughly 30 kbps over the same telephone lines, was not a useful bound for real systems. However, by the 1990s breakthroughs in hardware, modulation, and coding techniques brought commercial modems very close to the speeds predicted by Shannon in the 1940s. In fact, modems can exceed this 30 kbps limit on some telephone channels, but that is because transmission lines today are of better quality than in Shannon's day and thus have a higher received power than that used in his initial calculation. Advances in coding techniques over the last few decades include turbo codes [8], LDPC codes (originally developed by Gallager in the 1960s and rediscovered in the 1990s) [9], lattice codes [10], and polar codes [11]. All of these codes come within a fraction of a decibel of the Shannon capacity limit using practical blocklengths. These codes and their performance will be covered in Chapter 8.

Wireless channels typically exhibit flat or frequency-selective fading. In the next two sections we consider capacity of flat fading and frequency-selective fading channels under different assumptions regarding what is known about the channel.

Example 4.1: Consider a wireless channel where power falloff with distance follows the formula $P_r(d) = P_t(d_0/d)^3$ for $d_0 = 10\,\mathrm{m}$. Assume the channel has bandwidth $B = 30$ kHz and AWGN with noise PSD $N_0/2$,

where $N_0 = 10^{-9}$ W/Hz. For a transmit power of 1 W, find the capacity of this channel for a transmit-receive distance of 100 m and 1 km.

Solution: The received SNR is $\gamma = P_r(d)/(N_0 B) = .1^3/(10^{-9} \cdot 30 \cdot 10^3) = 33 = 15\,\text{dB}$ for $d = 100\,\text{m}$ and $\gamma = .01^3/(10^{-9} \cdot 30 \cdot 10^3) = .033 = -15\,\text{dB}$ for $d = 1000\,\text{m}$. The corresponding capacities are $C = B\log_2(1+\gamma) = 30000\log_2(1+33) = 152.6$ kbps for $d = 100\,\text{m}$ and $C = 30000\log_2(1+.033) = 1.4$ kbps for $d = 1000\,\text{m}$. Note the significant decrease in capacity at greater distances due to the path-loss exponent of 3, which greatly reduces received power as distance increases.

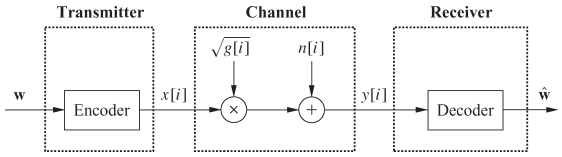


Figure 4.1: Flat fading channel and system model.

4.2 Capacity of Flat Fading Channels

4.2.1 Channel and System Model

We assume a discrete-time channel with stationary and ergodic time-varying gain $\sqrt{g[i]}$, $0 \le g[i]$, and AWGN $n[i]$, as shown in Figure 4.1. The channel power gain $g[i]$ follows a given distribution $p(g)$; for example, with Rayleigh fading $p(g)$ is exponential. We assume that $g[i]$ is independent of the channel input. The channel gain $g[i]$ can change at each time i, either as an independent and identically distributed (i.i.d.) process or with some correlation over time. In a *block fading channel*, $g[i]$ is constant over some blocklength T, after which time $g[i]$ changes to a new independent value based on the distribution $p(g)$. Let $\bar{P}$ denote the average transmit signal power, $N_0/2$ the noise PSD of $n[i]$, and B the received signal bandwidth. The instantaneous received SNR is then $\gamma[i] = \bar{P}g[i]/(N_0 B)$, $0 \le \gamma[i] < \infty$, and its expected value over all time is $\bar{\gamma} = \bar{P}\bar{g}/(N_0 B)$. Since $\bar{P}/(N_0 B)$ is a constant, the distribution of $g[i]$ determines the distribution of $\gamma[i]$ and vice versa.

The system model is also shown in Figure 4.1, where an input message $\mathbf{w}$ is sent from the transmitter to the receiver, which reconstructs an estimate $\hat{\mathbf{w}}$ of the transmitted message $\mathbf{w}$ from the received signal. The message is encoded into the codeword $\mathbf{x}$, which is transmitted over the time-varying channel as $x[i]$ at time i. The channel gain $g[i]$, also called the *channel side information* (CSI), changes during the transmission of the codeword.

The capacity of this channel depends on what is known about $g[i]$ at the transmitter and receiver. We will consider three different scenarios regarding this knowledge as follows.

1. *Channel distribution information* (CDI): The distribution of $g[i]$ is known to the transmitter and receiver.

2. *Receiver* CSI: The value of $g[i]$ is known to the receiver at time i, and both the transmitter and receiver know the distribution of $g[i]$.

3. *Transmitter and receiver CSI*: The value of $g[i]$ is known to the transmitter and receiver at time i, and both the transmitter and receiver know the distribution of $g[i]$.

Transmitter and receiver CSI allow the transmitter to adapt both its power and rate to the channel gain at time i, leading to the highest capacity of the three scenarios. Note that since the instantaneous SNR $\gamma[i]$ is just $g[i]$ multiplied by the constant $\bar{P}/(N_0B)$, known CSI or CDI about $g[i]$ yields the same information about $\gamma[i]$. Capacity for time-varying channels under assumptions other than these three are discussed in [12, 13].

4.2.2 Channel Distribution Information Known

We first consider the case where the channel gain distribution $p(g)$ or, equivalently, the distribution of SNR $p(\gamma)$ is known to the transmitter and receiver. For i.i.d. fading the capacity is given by (4.6), but solving for the capacity-achieving input distribution (i.e., the distribution achieving the maximum in that equation) can be quite complicated depending on the nature of the fading distribution. Moreover, fading correlation introduces channel memory, in which case the capacity-achieving input distribution is found by optimizing over input blocks, and this makes finding the solution even more difficult. For these reasons, finding the capacity-achieving input distribution and corresponding capacity of fading channels under CDI remains an open problem for almost all channel distributions.

The capacity-achieving input distribution and corresponding fading channel capacity under CDI are known for two specific models of interest: i.i.d. Rayleigh fading channels and finite-state Markov channels. In i.i.d. Rayleigh fading, the channel power gain is exponentially distributed and changes independently with each channel use. The optimal input distribution for this channel was shown in [14] to be discrete with a finite number of mass points, one of which is located at zero. This optimal distribution and its corresponding capacity must be found numerically. The lack of closed-form solutions for capacity or the optimal input distribution is somewhat surprising given that the fading follows the most common fading distribution and has no correlation structure. For flat fading channels that are not necessarily Rayleigh or i.i.d., upper and lower bounds on capacity have been determined in [15], and these bounds are tight at high SNRs.

Approximating Rayleigh fading channels via FSMCs was discussed in Chapter 3.2.4. This model approximates the fading correlation as a Markov process. Although the Markov nature of the fading dictates that the fading at a given time depends only on fading at the previous time sample, it turns out that the receiver must decode all past channel outputs jointly with the current output for optimal (i.e. capacity-achieving) decoding. This significantly complicates capacity analysis. The capacity of FSMCs has been derived for i.i.d. inputs in [16, 17] and for general inputs in [18]. Capacity of the FSMC depends on the limiting distribution of the channel conditioned on all past inputs and outputs, which can be computed recursively. As with the i.i.d. Rayleigh fading channel, the final result and complexity of the capacity analysis are high for this relatively simple fading model. This shows the difficulty of obtaining the capacity and related design insights on channels when only CDI is available.

4.2.3 Channel Side Information at Receiver

We now consider the case where the CSI $g[i]$ is known to the receiver at time i. Equivalently, $\gamma[i]$ is known to the receiver at time i. We also assume that both the transmitter and receiver know the distribution of $g[i]$. In this case there are two channel capacity definitions that are relevant to system design: Shannon capacity, also called *ergodic capacity*, and *capacity with outage*. As for the AWGN channel, Shannon capacity defines the maximum data rate that can be sent over the channel with asymptotically small error probability. Note that for Shannon capacity the rate transmitted over the channel is constant: the transmitter cannot adapt its transmission strategy relative to the CSI. Thus, poor channel states typically reduce Shannon capacity because the transmission strategy must incorporate the effect of these poor states. An alternate capacity definition for fading channels with receiver CSI is capacity with outage. This is defined as the maximum rate that can be transmitted over a channel with an outage probability corresponding to the probability that the transmission cannot be decoded with negligible error probability. The basic premise of capacity with outage is that a high data rate can be sent over the channel and

decoded correctly except when the channel is in a slow deep fade. By allowing the system to lose some data in the event of such deep fades, a higher data rate can be maintained than if all data must be received correctly regardless of the fading state, as is the case for Shannon capacity. The probability of outage characterizes the probability of data loss or, equivalently, of deep fading.

SHANNON (ERGODIC) CAPACITY

Shannon capacity of a fading channel with receiver CSI for an average power constraint $\bar{P}$ can be obtained from results in [19] as

$$C = \int_0^\infty B \log_2(1+\gamma) p(\gamma) d\gamma. \tag{4.7}$$

Note that this formula is a probabilistic average: the capacity C is equal to Shannon capacity for an AWGN channel with SNR γ, given by $B \log_2(1+\gamma)$ and averaged over the distribution of γ. That is why Shannon capacity is also called ergodic capacity. However, care must be taken in interpreting (4.7) as an average. In particular, it is incorrect to interpret (4.7) to mean that this average capacity is achieved by maintaining a capacity $B \log_2(1+\gamma)$ when the instantaneous SNR is γ, for only the receiver knows the instantaneous SNR $\gamma[i]$ and so the data rate transmitted over the channel is constant, regardless of γ. Note also that the capacity-achieving code must be sufficiently long that a received codeword is affected by all possible fading states. This can result in significant delay.

By Jensen's inequality,

$$\begin{aligned} \mathbf{E}[B \log_2(1+\gamma)] &= \int B \log_2(1+\gamma) p(\gamma) d\gamma \\ &\leq B \log_2(1+\mathbf{E}[\gamma]) = B \log_2(1+\bar{\gamma}), \end{aligned} \tag{4.8}$$

where $\bar{\gamma}$ is the average SNR on the channel. Thus we see that the Shannon capacity of a fading channel with receiver CSI only is less than the Shannon capacity of an AWGN channel with the same average SNR. In other words, fading reduces Shannon capacity when only the receiver has CSI. Moreover, without transmitter CSI, the code design must incorporate the channel correlation statistics, and the complexity of the maximum likelihood decoder will be proportional to the channel decorrelation time. In addition, if the receiver CSI is not perfect, capacity can be significantly decreased [20].

Example 4.2: Consider a flat fading channel with i.i.d. channel gain $\sqrt{g[i]}$, which can take on three possible values: $\sqrt{g_1} = .05$ with probability $p_1 = .1$, $\sqrt{g_2} = .5$ with probability $p_2 = .5$, and $\sqrt{g_3} = 1$ with probability $p_3 = .4$. The transmit power is 10 mW, the noise power spectral density $N_0/2$ has $N_0 = 10^{-9}$ W/Hz, and the channel bandwidth is 30 kHz. Assume the receiver has knowledge of the instantaneous value of $g[i]$ but the transmitter does not. Find the Shannon capacity of this channel and compare with the capacity of an AWGN channel with the same average SNR.

Solution: The channel has three possible received SNRs: $\gamma_1 = P_t g_1/(N_0 B) = .01 \cdot (.05^2)/(30000 \cdot 10^{-9}) = .8333 = -.79\,\text{dB}$, $\gamma_2 = P_t g_2/(N_0 B) = .01 \cdot (.5^2)/(30000 \cdot 10^{-9}) = 83.333 = 19.2\,\text{dB}$, and $\gamma_3 = P_t g_3/(N_0 B) = .01/(30000 \cdot 10^{-9}) = 333.33 = 25\,\text{dB}$. The probabilities associated with each of these SNR values are $p(\gamma_1) = .1$, $p(\gamma_2) = .5$, and $p(\gamma_3) = .4$. Thus, the Shannon capacity is given by

$$
\begin{aligned}
C &= \sum_i B\log_2(1+\gamma_i)p(\gamma_i) \\
&= 30000(.1\log_2(1.8333) + .5\log_2(84.333) + .4\log_2(334.33)) \\
&= 199.26 \text{ kbps.}
\end{aligned} \tag{4.9}
$$

The average SNR for this channel is $\bar{\gamma} = .1(.8333) + .5(83.33) + .4(333.33) = 175.08 = 22.43$ dB. The capacity of an AWGN channel with this SNR is $C = B\log_2(1 + 175.08) = 223.8$ kbps. Note that this rate is about 25 kbps larger than that of the flat fading channel with receiver CSI and the same average SNR.

CAPACITY WITH OUTAGE

Capacity with outage applies to slowly varying channels, where the instantaneous SNR γ is constant over a large number of transmissions (a transmission burst) and then changes to a new value based on the fading distribution. With this model, if the channel has received SNR γ during a burst then data can be sent over the channel at rate $B\log_2(1+\gamma)$ with negligible probability of error.[1] Since the transmitter does not know the SNR value γ, it must fix a transmission rate independent of the instantaneous received SNR.

Capacity with outage allows bits sent over a given transmission burst to be decoded at the end of the burst with some probability that these bits will be decoded incorrectly. Specifically, the transmitter fixes a minimum received SNR $\gamma_{\min}$ and encodes for a data rate $C = B\log_2(1+\gamma_{\min})$. The data is correctly received if the instantaneous received SNR is greater than or equal to $\gamma_{\min}$ [21, 22]. If the received SNR is below $\gamma_{\min}$ then the bits received over that transmission burst cannot be decoded correctly with probability approaching 1, and the receiver declares an outage. The probability of outage is thus $P_{\text{out}} = p(\gamma < \gamma_{\min})$. The average rate correctly received over many transmission bursts is $C_{\text{out}} = (1 - P_{\text{out}})B\log_2(1+\gamma_{\min})$ since data is only correctly received on $1 - P_{\text{out}}$ transmissions. The value of $\gamma_{\min}$ is a design parameter based on the acceptable outage probability. Capacity with outage is typically characterized by a plot of capacity versus outage, as shown in Figure 4.2. In this figure we plot the normalized capacity $C/B = \log_2(1+\gamma_{\min})$ as a function of outage probability $P_{\text{out}} = p(\gamma < \gamma_{\min})$ for a Rayleigh fading channel (γ exponentially distributed) with $\bar{\gamma} = 20$ dB. We see that capacity approaches zero for small outage probability, due to the requirement that bits transmitted under severe fading must be decoded correctly, and increases dramatically as outage probability increases. Note, however, that these high capacity values for large outage probabilities have higher probability of incorrect data reception. The average rate correctly received can be maximized by finding the $\gamma_{\min}$ (or, equivalently, the P_{out}) that maximizes C_{out}.

Example 4.3: Assume the same channel as in the previous example, with a bandwidth of 30 kHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Find the capacity versus outage for this channel, and find the average rate correctly received for outage probabilities $P_{\text{out}} < .1$, $P_{\text{out}} = .1$, and $P_{\text{out}} = .6$.

Solution: For time-varying channels with discrete SNR values, the capacity versus outage is a staircase function. Specifically, for $P_{\text{out}} < .1$ we must decode correctly in all channel states. The minimum received SNR for P_{out} in this range of values is that of the weakest channel: $\gamma_{\min} = \gamma_1$, and the corresponding capacity is $C = B\log_2(1+\gamma_{\min}) = 30000\log_2(1.833) = 26.23$ kbps. For $.1 \le P_{\text{out}} < .6$ we can decode incorrectly when the channel is in the weakest state only. Then $\gamma_{\min} = \gamma_2$ and the corresponding capacity is $C = B\log_2(1+\gamma_{\min}) =$

[1] The assumption of constant fading over a large number of transmissions is needed because codes that achieve capacity require very large blocklengths.

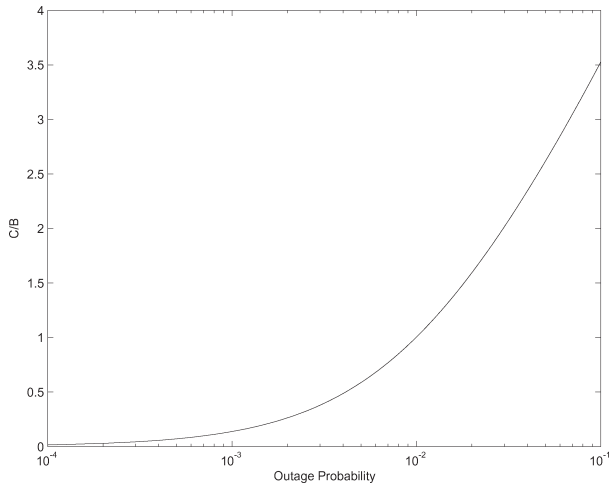


Figure 4.2: Normalized capacity (C/B) versus outage probability.

$30000 \log_2(84.33) = 191.94$ kbps. For $.6 \leq P_{\text{out}} < 1$ we can decode incorrectly if the channel has received SNR γ_1 or γ_2. Then $\gamma_{\min} = \gamma_3$ and the corresponding capacity is $C = B \log_2(1 + \gamma_{\min}) = 30000 \log_2(334.33) = 251.55$ kbps. Thus, capacity versus outage has $C = 26.23$ kbps for $P_{\text{out}} < .1$, $C = 191.94$ kbps for $.1 \leq P_{\text{out}} < .6$, and $C = 251.55$ kbps for $.6 \leq P_{\text{out}} < 1$.

For $P_{\text{out}} < .1$, data transmitted at rates close to capacity $C = 26.23$ kbps are always correctly received because the channel can always support this data rate. For $P_{\text{out}} = .1$ we transmit at rates close to $C = 191.94$ kbps, but we can correctly decode these data only when the SNR is γ_2 or γ_3, so the rate correctly received is $(1 - .1)191940 = 172.75$ kbps. For $P_{\text{out}} = .6$ we transmit at rates close to $C = 251.55$ kbps but we can correctly decode these data only when the SNR is γ_3, so the rate correctly received is $(1 - .6)251550 = 100.62$ kbps. It is likely that a good engineering design for this channel would send data at a rate close to 191.94 kbps, since it would be received incorrectly at most 10% of this time and the data rate would be almost an order of magnitude higher than sending at a rate commensurate with the worst-case channel capacity. However, 10% retransmission probability is too high for some applications, in which case the system would be designed for the 26.23 kbps data rate with no retransmissions. Design issues regarding acceptable retransmission probability are discussed in Chapter 14.

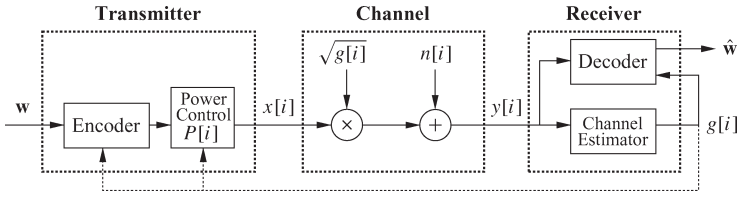


Figure 4.3: System model with transmitter and receiver CSI.

4.2.4 Channel Side Information at Transmitter and Receiver

If both the transmitter and receiver have CSI then the transmitter can adapt its transmission strategy relative to this CSI, as shown in Figure 4.3. In this case there is no notion of capacity versus outage where the transmitter sends

bits that cannot be decoded, since the transmitter knows the channel and thus will not send bits unless they can be decoded correctly. In this section we will derive Shannon capacity assuming optimal power and rate adaptation relative to the CSI; we also introduce alternate capacity definitions and their power and rate adaptation strategies.

SHANNON CAPACITY

We now consider the Shannon capacity when the channel power gain $g[i]$ is known to both the transmitter and receiver at time i. The Shannon capacity of a time-varying channel with side information about the channel state at both the transmitter and receiver was originally considered by Wolfowitz for the following model. Let $s[i]$ be a stationary and ergodic stochastic process representing the channel state, which takes values on a finite set S of discrete memoryless channels. Let C_s denote the capacity of a particular channel $s \in S$ and let $p(s)$ denote the probability, or fraction of time, that the channel is in state s. The capacity of this time-varying channel is then given by Theorem 4.6.1 of [23]:

$$C = \sum_{s \in \mathcal{S}} C_s p(s). \tag{4.10}$$

We now apply this formula to the system model in Figure 4.3. We know that the capacity of an AWGN channel with average received SNR γ is $C_\gamma = B \log_2(1+\gamma)$. Let $p(\gamma) = p(\gamma[i] = \gamma)$ denote the distribution of the received SNR. By (4.10), the capacity of the fading channel with transmitter and receiver side information is thus[2]

$$C = \int_0^\infty C_\gamma p(\gamma) d\gamma = \int_0^\infty B \log_2(1+\gamma) p(\gamma) d\gamma. \tag{4.11}$$

We see that, without power adaptation, (4.7) and (4.11) are the same, so transmitter side information does not increase capacity unless power is also adapted.

Now let us allow the transmit power $P(\gamma)$ to vary with γ subject to an average power constraint $\bar{P}$:

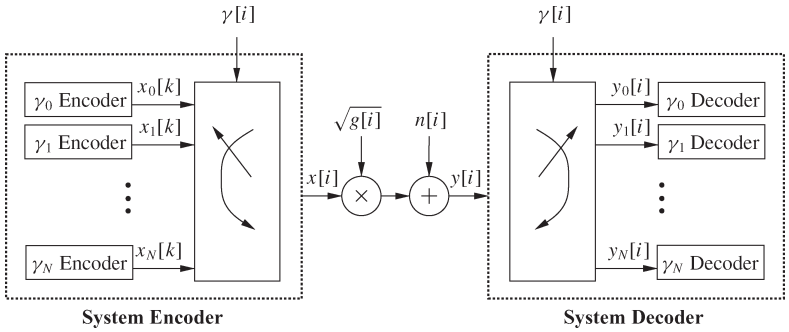


Figure 4.4: Multiplexed coding and decoding.

$$\int_0^\infty P(\gamma) p(\gamma) d\gamma \leq \bar{P}. \tag{4.12}$$

With this additional constraint, we cannot apply (4.11) directly to obtain the capacity. However, we expect that the capacity with this average power constraint will be the average capacity given by (4.11) with the power optimally distributed over time. This motivates our definition of the fading channel capacity with average power constraint (4.12) as

[2]Wolfowitz's result was for γ ranging over a finite set, but it can be extended to infinite sets [24].

$$C = \max_{P(\gamma):\int P(\gamma)p(\gamma)d\gamma=\bar{P}} \int_0^\infty B\log_2\left(1+\frac{P(\gamma)\gamma}{\bar{P}}\right)p(\gamma)d\gamma. \tag{4.13}$$

It is proved in [24] that the capacity given in (4.13) can be achieved and that any rate larger than this capacity has probability of error bounded away from zero. The main idea behind the proof is a "time diversity" system with multiplexed input and demultiplexed output, as shown in Figure 4.4. Specifically, we first quantize the range of fading values to a finite set $\{\gamma_j : 1 \le j \le N\}$. For each γ_j, we design an encoder–decoder pair for an AWGN channel with SNR γ_j. The input x_j for encoder γ_j has average power $P(\gamma_j)$ and data rate $R_j = C_j$, where C_j is the capacity of a time-invariant AWGN channel with received SNR $P(\gamma_j)\gamma_j/\bar{P}$. These encoder–decoder pairs correspond to a set of input and output ports associated with each γ_j. When $\gamma[i] \approx \gamma_j$, the corresponding pair of ports are connected through the channel. The codewords associated with each γ_j are thus multiplexed together for transmission and then demultiplexed at the channel output. This effectively reduces the time-varying channel to a set of time-invariant channels in parallel, where the jth channel operates only when $\gamma[i] \approx \gamma_j$. The average rate on the channel is just the sum of rates associated with each of the γ_j channels weighted by $p(\gamma_j)$, the percentage of time that the channel SNR equals γ_j. This yields the average capacity formula (4.13).

To find the optimal power allocation $P(\gamma)$, we form the Lagrangian

$$J(P(\gamma)) = \int_0^\infty B\log_2\left(1+\frac{\gamma P(\gamma)}{\bar{P}}\right)p(\gamma)d\gamma - \lambda\int_0^\infty P(\gamma)p(\gamma)d\gamma. \tag{4.14}$$

Next we differentiate the Lagrangian and set the derivative equal to zero:

$$\frac{\partial J(P(\gamma))}{\partial P(\gamma)} = \left[\left(\frac{B/\ln(2)}{1+\gamma P(\gamma)/\bar{P}}\right)\frac{\gamma}{\bar{P}} - \lambda\right]p(\gamma) = 0. \tag{4.15}$$

Solving for $P(\gamma)$ with the constraint that $P(\gamma) > 0$ yields the optimal power adaptation that maximizes (4.13) as

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} 1/\gamma_0 - 1/\gamma & \gamma \ge \gamma_0, \\ 0 & \gamma < \gamma_0, \end{cases} \tag{4.16}$$

for some "cutoff" value γ_0. If $\gamma[i]$ is below this cutoff then no data is transmitted over the ith time interval, so the channel is used at time i only if $\gamma_0 \le \gamma[i] < \infty$. Substituting (4.16) into (4.13) then yields the capacity formula:

$$C = \int_{\gamma_0}^\infty B\log_2\left(\frac{\gamma}{\gamma_0}\right)p(\gamma)d\gamma. \tag{4.17}$$

The multiplexing nature of the capacity-achieving coding strategy indicates that (4.17) is achieved with a time-varying data rate, where the rate corresponding to the instantaneous SNR γ is $B\log_2(\gamma/\gamma_0)$. Since γ_0 is constant, this means that as the instantaneous SNR increases, the data rate sent over the channel for that instantaneous SNR also increases. Note that this multiplexing strategy is not the only way to achieve capacity (4.17): it can also be achieved by adapting the transmit power and sending at a fixed rate [25]. We will see in Section 4.2.6 that for Rayleigh fading this capacity can exceed that of an AWGN channel with the same average SNR – in contrast to the case of receiver CSI only, where fading always decreases capacity.

Note that the optimal power allocation policy (4.16) depends on the fading distribution $p(\gamma)$ only through the cutoff value γ_0. This cutoff value is found from the power constraint. Specifically, rearranging the power constraint (4.12) and replacing the inequality with equality (since using the maximum available power will always be optimal) yields the power constraint

$$\int_0^\infty \frac{P(\gamma)}{\bar{P}}p(\gamma)d\gamma = 1. \tag{4.18}$$

If we now substitute the optimal power adaptation (4.16) into this expression then the cutoff value γ_0 must satisfy

$$\int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma} \right) p(\gamma) d\gamma = 1. \tag{4.19}$$

Observe that this expression depends only on the distribution $p(\gamma)$. The value for γ_0 must be found numerically [26] because no closed-form solutions exist for typical continuous distributions $p(\gamma)$.

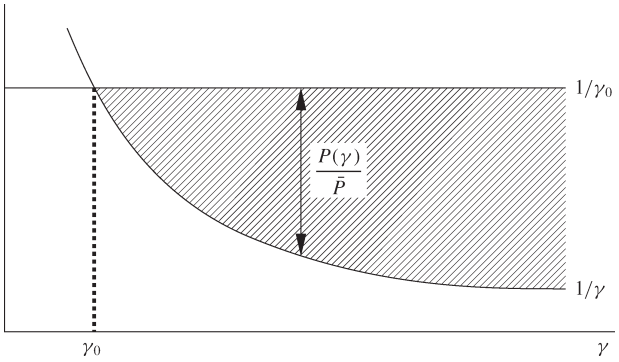


Figure 4.5: Optimal power allocation: water-filling.

Since γ is time varying, the maximizing power adaptation policy of (4.16) is a water filling formula in time, as illustrated in Figure 4.5. This curve shows how much power is allocated to the channel for instantaneous SNR $\gamma(t) = \gamma$. The water-filling terminology refers to the fact that the line $1/\gamma$ sketches out the bottom of a bowl, and power is poured into the bowl to a constant water level of $1/\gamma_0$. The amount of power allocated for a given γ equals $1/\gamma_0 - 1/\gamma$, the amount of water between the bottom of the bowl $(1/\gamma)$ and the constant water line $(1/\gamma_0)$. The intuition behind water-filling is to take advantage of good channel conditions: when channel conditions are good (γ large), more power and a higher data rate are sent over the channel. As channel quality degrades (γ small), less power and rate are sent over the channel. If the instantaneous SNR falls below the cutoff value, the channel is not used. Adaptive modulation and coding techniques that follow this principle were developed in [27, 28] and are discussed in Chapter 9.

Note that the multiplexing argument sketching how capacity (4.13) is achieved applies to any power adaptation policy. That is, for any power adaptation policy $P(\gamma)$ with average power $\bar{P}$, the capacity

$$C = \int_0^{\infty} B \log_2 \left(1 + \frac{P(\gamma)\gamma}{\bar{P}} \right) p(\gamma) d\gamma \tag{4.20}$$

can be achieved with arbitrarily small error probability. Of course this capacity cannot exceed (4.13), where power adaptation is optimized to maximize capacity. However, there are scenarios where a suboptimal power adaptation policy might have desirable properties that outweigh capacity maximization. In the next two sections we discuss two such suboptimal policies, which result in constant data rate systems, in contrast to the variable-rate transmission policy that achieves the capacity in (4.13).

Example 4.4: Assume the same channel as in the previous example, with a bandwidth of 30 kHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Find the ergodic capacity of this channel assuming that both transmitter and receiver have instantaneous CSI.
Solution: We know the optimal power allocation is water-filling, and we need to find the cutoff value γ_0 that

satisfies the discrete version of (4.19) given by

$$\sum_{\gamma_i \geq \gamma_0} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_i} \right) p(\gamma_i) = 1. \tag{4.21}$$

We first assume that all channel states are used to obtain γ_0 (i.e., we assume $\gamma_0 \leq \min_i \gamma_i$) and see if the resulting cutoff value is below that of the weakest channel. If not then we have an inconsistency, and must redo the calculation assuming at least one of the channel states is not used. Applying the discrete power allocation formula to our channel model yields

$$\sum_{i=1}^{3} \frac{p(\gamma_i)}{\gamma_0} - \sum_{i=1}^{3} \frac{p(\gamma_i)}{\gamma_i} = 1 \quad \Rightarrow \quad \frac{1}{\gamma_0} = 1 + \sum_{i=1}^{3} \frac{p(\gamma_i)}{\gamma_i} = 1 + \left(\frac{.1}{.8333} + \frac{.5}{83.33} + \frac{.4}{333.33} \right) = 1.13. \tag{4.22}$$

Solving for γ_0 yields $\gamma_0 = 1/1.13 = .89 > .8333 = \gamma_1$. Since this value of γ_0 is greater than the SNR in the weakest channel, this result is inconsistent because the channel should only be used for SNRs above the cutoff value. Therefore, we now redo the calculation assuming that the weakest state is not used. Then the discrete power allocation becomes

$$\sum_{i=2}^{3} \frac{p(\gamma_i)}{\gamma_0} - \sum_{i=2}^{3} \frac{p(\gamma_i)}{\gamma_i} = 1 \quad \Rightarrow \quad \frac{.9}{\gamma_0} = 1 + \sum_{i=2}^{3} \frac{p(\gamma_i)}{\gamma_i} = 1 + \left(\frac{.5}{83.33} + \frac{.4}{333.33} \right) = 1.0072. \tag{4.23}$$

Solving for γ_0 yields $\gamma_0 = .89$. Hence, by assuming that the weakest channel with SNR γ_1 is not used, we obtain a consistent value for γ_0 with $\gamma_1 < \gamma_0 \leq \gamma_2$. The capacity of the channel then becomes

$$\begin{aligned} C &= \sum_{i=2}^{3} B \log_2 \left(\frac{\gamma_i}{\gamma_0} \right) p(\gamma_i) \\ &= 30000 \left(.5 \log_2 \frac{83.33}{.89} + .4 \log_2 \frac{333.33}{.89} \right) = 200.82 \text{ kbps}. \end{aligned} \tag{4.24}$$

Comparing with the results of Example 4.3 we see that this rate is only slightly higher than for the case of receiver CSI only, and it is still significantly below that of an AWGN channel with the same average SNR. This is because the average SNR for the channel in this example is relatively high: for low-SNR channels, capacity with flat fading can exceed that of the AWGN channel with the same average SNR if we take advantage of the rare times when the fading channel is in a very good state.

ZERO-OUTAGE CAPACITY AND CHANNEL INVERSION

We now consider a suboptimal transmitter adaptation scheme where the transmitter uses the CSI to maintain a constant received power; that is, it inverts the channel fading. The channel then appears to the encoder and decoder as a time-invariant AWGN channel. This power adaptation, called *channel inversion*, is given by $P(\gamma)/\bar{P} = \sigma/\gamma$, where σ equals the constant received SNR that can be maintained with the transmit power constraint (4.12). The constant σ thus satisfies $\int (\sigma/\gamma) p(\gamma) d\gamma = 1$, so $\sigma = 1/\mathbf{E}[1/\gamma]$.

Fading channel capacity with channel inversion is just the capacity of an AWGN channel with SNR σ:

$$C = B\log_2[1+\sigma] = B\log_2\left[1+\frac{1}{\mathbf{E}[1/\gamma]}\right]. \tag{4.25}$$

The capacity-achieving transmission strategy for this capacity uses a fixed-rate encoder and decoder designed for an AWGN channel with SNR σ. This has the advantage of maintaining a fixed data rate over the channel regardless of channel conditions. For this reason the channel capacity given in (4.25) is called *zero-outage capacity*, since the data rate is fixed under all channel conditions and there is no channel outage. Note that there exist practical coding techniques that achieve near-capacity data rates on AWGN channels, so the zero-outage capacity can be approximately achieved in practice.

Zero-outage capacity can exhibit a large data-rate reduction relative to Shannon capacity in extreme fading environments. In Rayleigh fading, for example, $\mathbf{E}[1/\gamma]$ is infinite and thus the zero-outage capacity given by (4.25) is zero. Channel inversion is common in spread-spectrum systems with near–far interference imbalances [29]. It is also the simplest scheme to implement because the encoder and decoder are designed for an AWGN channel, independent of the fading statistics.

Example 4.5: Assume the same channel as in the previous example, with a bandwidth of 30 kHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Assuming transmitter and receiver CSI, find the zero-outage capacity of this channel.
Solution: The zero-outage capacity is $C = B\log_2[1+\sigma]$, where $\sigma = 1/\mathbf{E}[1/\gamma]$. Since

$$\mathbf{E}\left[\frac{1}{\gamma}\right] = \frac{.1}{.8333} + \frac{.5}{83.33} + \frac{.4}{333.33} = .1272, \tag{4.26}$$

we have $C = 30000\log_2[1+1/.1272] = 94.43$ kbps. Note that this is less than half of the Shannon capacity with optimal water-filling adaptation.

OUTAGE CAPACITY AND TRUNCATED CHANNEL INVERSION

The reason that zero-outage capacity may be significantly smaller than Shannon capacity on a fading channel is the requirement of maintaining a constant data rate in all fading states. By suspending transmission in particularly bad fading states (outage channel states), we can maintain a higher constant data rate in the other states and thereby significantly increase capacity. The *outage capacity* is defined as the maximum data rate that can be maintained in all non-outage channel states multiplied by the probability of non-outage. Outage capacity is achieved with a *truncated channel inversion* policy for power adaptation that compensates for fading only above a certain cutoff fade depth γ_0:

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \sigma/\gamma & \gamma \geq \gamma_0, \\ 0 & \gamma < \gamma_0, \end{cases} \tag{4.27}$$

where γ_0 is based on the outage probability: $P_{\text{out}} = p(\gamma < \gamma_0)$. Since the channel is only used when $\gamma \geq \gamma_0$, the power constraint (4.12) yields $\sigma = 1/\mathbf{E}_{\gamma_0}[1/\gamma]$, where

$$\mathbf{E}_{\gamma_0}\left[\frac{1}{\gamma}\right] \triangleq \int_{\gamma_0}^{\infty} \frac{1}{\gamma} p(\gamma) d\gamma. \tag{4.28}$$

The outage capacity associated with a given outage probability P_{out} and corresponding cutoff γ_0 is given by

$$C(P_{\text{out}}) = B\log_2\left(1 + \frac{1}{\mathbf{E}_{\gamma_0}[1/\gamma]}\right)p(\gamma \geq \gamma_0). \tag{4.29}$$

We can also obtain the *maximum outage capacity* by maximizing outage capacity over all possible γ_0:

$$C = \max_{\gamma_0} B\log_2\left(1 + \frac{1}{\mathbf{E}_{\gamma_0}[1/\gamma]}\right)p(\gamma \geq \gamma_0). \tag{4.30}$$

This maximum outage capacity will still be less than Shannon capacity (4.17) because truncated channel inversion is a suboptimal transmission strategy. However, the transmit and receive strategies associated with inversion or truncated inversion may be easier to implement or have lower complexity than the water-filling schemes associated with Shannon capacity.

Example 4.6: Assume the same channel as in the previous example, with a bandwidth of 30 kHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Find the outage capacity of this channel and associated outage probabilities for cutoff values $\gamma_0 = .84$ and $\gamma_0 = 83.4$. Which of these cutoff values yields a larger outage capacity?
Solution: For $\gamma_0 = .84$ we use the channel when the SNR is γ_2 or γ_3, so $\mathbf{E}_{\gamma_0}[1/\gamma] = \sum_{i=2}^{3} p(\gamma_i)/\gamma_i = .5/83.33 + .4/333.33 = .0072$. The outage capacity is $C = B\log_2(1 + 1/\mathbf{E}_{\gamma_0}[1/\gamma])p(\gamma \geq \gamma_0) = 30000\log_2(1 + 138.88) \cdot .9 = 192.457$ kbps. For $\gamma_0 = 83.34$ we use the channel when the SNR is γ_3 only, so $\mathbf{E}_{\gamma_0}[1/\gamma] = p(\gamma_3)/\gamma_3 = .4/333.33 = .0012$. The capacity is $C = B\log_2(1 + 1/\mathbf{E}_{\gamma_0}[1/\gamma])p(\gamma \geq \gamma_0) = 30000\log_2(1 + 833.33) \cdot .4 = 116.45$ kbps. The outage capacity is larger when the channel is used for SNRs γ_2 and γ_3. Even though the SNR γ_3 is significantly larger than γ_2, the fact that this larger SNR occurs only 40% of the time makes it inefficient to only use the channel in this best state.

4.2.5 Capacity with Receiver Diversity

Receiver diversity is a well-known technique for improving the performance of wireless communications in fading channels. The main advantage of receiver diversity is that it mitigates the fluctuations due to fading so that the channel appears more like an AWGN channel. More details on receiver diversity and its performance will be given in Chapter 7. Since receiver diversity mitigates the impact of fading, an interesting question is whether it also increases the capacity of a fading channel. The capacity calculation under diversity combining requires first that the distribution of the received SNR $p(\gamma)$ under the given diversity combining technique be obtained. Once this distribution is known, it can be substituted into any of the capacity formulas already given to obtain the capacity under diversity combining. The specific capacity formula used depends on the assumptions about channel side information; for example, in the case of perfect transmitter and receiver CSI the formula (4.17) would be used. For different CSI assumptions, capacity (under both maximal ratio and selection combining diversity) was computed in [26]. It was found that, as expected, the capacity with perfect transmitter and receiver CSI is greater than with receiver CSI only, which in turn is greater than with channel inversion. The performance gap of these different formulas decreases as the number of antenna branches increases. This trend is expected because a large number of antenna branches makes the channel look like an AWGN channel, for which all of the different capacity formulas have roughly the same performance.

As originally derived in [30, 31, 32], the capacity of a fading channel with multiple antennas at the transmitter and receiver (a multiple input and output or MIMO channel) is M times larger than the channel capacity without multiple antennas, where $M = \min(M_t,\ M_r)$ for M_t the number of transmit antennas and M_r the number of

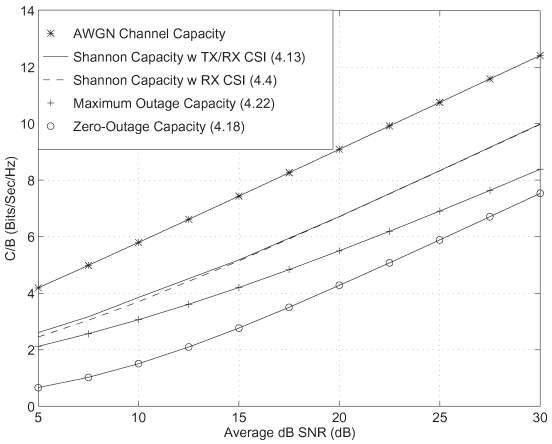


Figure 4.6: Capacity in log-normal fading.

receive antennas. Capacity of MIMO channels under different fading and channel CSI assumptions is discussed in Chapter 10.

4.2.6 Capacity Comparisons

In this section we compare capacity with transmitter and receiver CSI for different power allocation policies along with the capacity under receiver CSI only. Figures 4.6, 4.7, and 4.8 show plots of the different capacities (4.7), (4.17), (4.25), and (4.30) as a function of average received SNR for log-normal fading ($\sigma = 8\,\mathrm{dB}$ standard deviation), Rayleigh fading, and Nakagami fading (with Nakagami parameter $m = 2$). Nakagami fading with $m = 2$ is roughly equivalent to Rayleigh fading with two-antenna receiver diversity. The capacity in AWGN for the same average power is also shown for comparison. Note that the capacity in log-normal fading is plotted relative to average dB SNR (μ_{dB}), not average SNR in dB ($10\log_{10}\mu$): the relation between these values, as given by (2.45), is $10\log_{10}\mu = \mu_{\mathrm{dB}} + \sigma^2_{\mathrm{dB}}\ln(10)/20$.

Several observations in this comparison are worth noting. First, the figures show that the capacity of the AWGN channel is larger than that of the fading channel for all cases. However, at low SNRs the AWGN and fading channel with transmitter and receiver CSI have almost the same capacity. In fact, at low SNRs (below 0 dB), capacity of the fading channel with transmitter and receiver CSI is larger than the corresponding AWGN channel capacity. That is because the AWGN channel always has the same low SNR, thereby limiting its capacity. A fading channel with this same low average SNR will occasionally have a high SNR, since the distribution has infinite range. Thus, if high power and rate are transmitted over the channel during these very infrequent large SNR values, the capacity will be greater than on the AWGN channel with the same low average SNR.

The severity of the fading is indicated by the Nakagami parameter m, where $m = 1$ for Rayleigh fading and $m = \infty$ for an AWGN channel without fading. Thus, comparing Figures 4.7 and 4.8 we see that, as the severity of the fading decreases (Rayleigh to Nakagami with $m = 2$), the capacity difference between the various adaptive policies also decreases, and their respective capacities approach that of the AWGN channel.

The difference between the capacity curves under transmitter and receiver CSI (4.17) and receiver CSI only (4.7) are negligible in all cases. Recalling that capacity under receiver CSI only (4.7) and under transmitter and receiver CSI without power adaptation (4.11) are the same, we conclude that, if the transmission rate is adapted relative to the channel, then adapting the power as well yields a negligible capacity gain. It also indicates that transmitter adaptation yields a negligible capacity gain relative to using only receiver side information. In severe fading conditions (Rayleigh and log-normal fading), maximum outage capacity exhibits a 1–5dB rate penalty

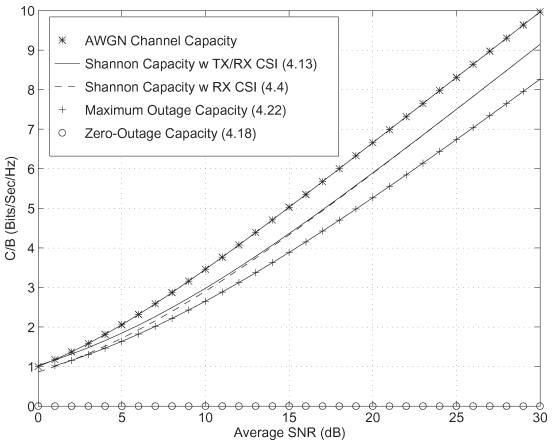


Figure 4.7: Capacity in Rayleigh fading.

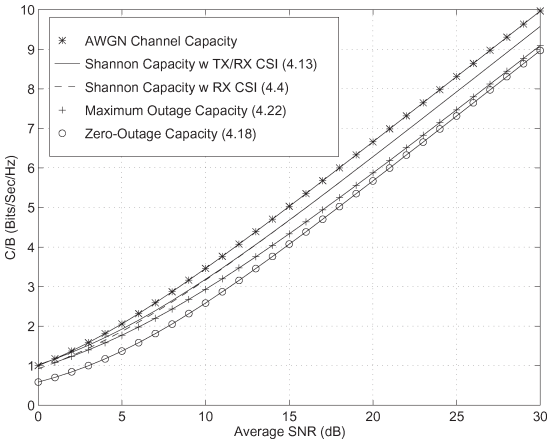


Figure 4.8: Capacity in Nakagami fading ($m = 2$).

and zero-outage capacity yields a large capacity loss relative to Shannon capacity. However, under mild fading conditions (Nakagami with $m = 2$) the Shannon, maximum outage, and zero-outage capacities are within 3 dB of each other and within 4 dB of the AWGN channel capacity. These differences will further decrease as the fading diminishes ($m \to \infty$ for Nakagami fading).

We can view these results as a trade-off between capacity and complexity. The adaptive policy with transmitter and receiver side information requires more complexity in the transmitter (and typically also requires a feedback path between the receiver and transmitter to obtain the side information). However, the decoder in the receiver is relatively simple. The nonadaptive policy has a relatively simple transmission scheme, but its code design must use the channel correlation statistics (often unknown) and the decoder complexity is proportional to the channel decorrelation time. The channel inversion and truncated inversion policies use codes designed for AWGN channels and thus are the least complex to implement, but in severe fading conditions they exhibit large capacity losses relative to the other techniques.

In general, Shannon capacity analysis does not show how to design adaptive or non-adaptive techniques for real systems. Achievable rates for adaptive trellis-coded MQAM have been investigated in [28], where a simple four-state trellis code combined with adaptive six-constellation MQAM modulation was shown to achieve rates within 7 dB of the Shannon capacity (4.13) in Figures 4.6 and 4.7. More complex codes further close the gap to

the Shannon limit of fading channels with transmitter adaptation.

4.3 Capacity of Frequency-Selective Fading Channels

In this section we examine the Shannon capacity of frequency-selective fading channels. We first consider the capacity of a time-invariant frequency-selective fading channel. This capacity analysis is like that of a flat fading channel but with the time axis replaced by the frequency axis. Then we discuss the capacity of time-varying frequency-selective fading channels.

4.3.1 Time-Invariant Channels

Consider a time-invariant channel with frequency response $H(f)$, as shown in Figure 4.9. Assume a total transmit power constraint P. When the channel is time invariant it is typically assumed that $H(f)$ is known to both the transmitter and receiver. The capacity of time-invariant channels under different assumptions about channel knowledge is discussed in [23, 25].

Let us first assume that $H(f)$ is *block fading*, so that frequency is divided into subchannels of bandwidth B with $H(f) = H_j$ constant over each subchannel, as shown in Figure 4.10.

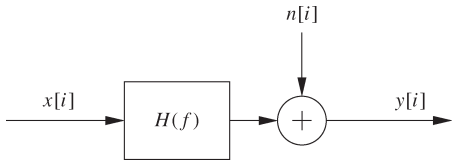


Figure 4.9: Time-invariant frequency-selective fading channel.

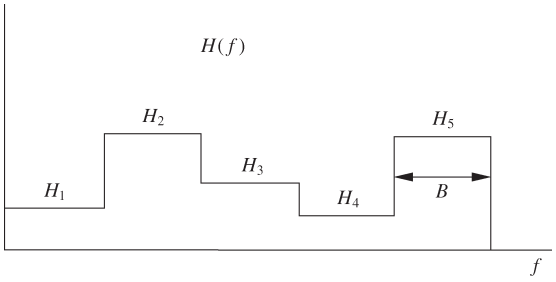


Figure 4.10: Block frequency-selective fading.

The frequency-selective fading channel thus consists of a set of AWGN channels in parallel with SNR $|H_j|^2 P_j/(N_0 B)$ on the jth channel, where P_j is the power allocated to the jth channel in this parallel set subject to the power constraint $\sum_j P_j \leq P$.

The capacity of this parallel set of channels is the sum of rates on each channel with power optimally allocated over all channels [5, 6]:

$$C = \sum_{\max P_j : \sum_j P_j \leq P} B \log_2 \left(1 + \frac{|H_j|^2 P_j}{N_0 B}\right). \tag{4.31}$$

Note that this is similar to the capacity and optimal power allocation for a flat fading channel, with power and rate changing over frequency in a deterministic way rather than over time in a probabilistic way. The optimal power

allocation is found via the same Lagrangian technique used in the flat-fading case, which leads to the water-filling power allocation

$$\frac{P_j}{P} = \begin{cases} 1/\gamma_0 - 1/\gamma_j & \gamma_j \geq \gamma_0, \\ 0 & \gamma_j < \gamma_0, \end{cases} \tag{4.32}$$

for some cutoff value γ_0, where $\gamma_j = |H_j|^2 P/(N_0 B)$ is the SNR associated with the jth channel assuming it is allocated the entire power budget P.[3] This optimal power allocation is illustrated in Figure 4.11. The cutoff value is obtained by substituting the power adaptation formula into the power constraint, so γ_0 must satisfy

$$\sum_{j:\gamma_j \geq \gamma_0} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_j} \right) = 1. \tag{4.33}$$

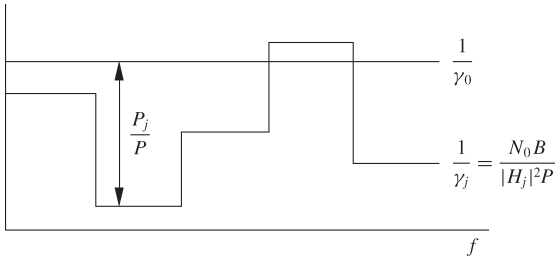


Figure 4.11: Water-filling in frequency-selective block fading.

The capacity then becomes

$$C = \sum_{j:\gamma_j \geq \gamma_0} B \log_2 \left(\frac{\gamma_j}{\gamma_0} \right). \tag{4.34}$$

This capacity is achieved by transmitting at different rates and powers over each subchannel. Multicarrier modulation uses the same technique in adaptive loading, as discussed in more detail in Section 12.3.

When $H(f)$ is continuous, the capacity under power constraint P is similar to the case of the block fading channel, with some mathematical intricacies needed to show that the channel capacity is given by

$$C = \max_{P(f): \int P(f) df \leq P} \int \log_2 \left(1 + \frac{|H(f)|^2 P(f)}{N_0} \right) df. \tag{4.35}$$

The expression inside the integral can be thought of as the incremental capacity associated with given frequency f over the bandwidth df with power allocation $P(f)$ and channel gain $|H(f)|^2$. This result is formally proven using a Karhunen-Loeve expansion of the channel $h(t)$ to create an equivalent set of parallel independent channels [5, Chap. 8.5]. An alternate proof [33] decomposes the channel into a parallel set using the discrete Fourier transform (DFT); the same premise is used in the discrete implementation of multicarrier modulation described in Section 12.4.

The power allocation over frequency, $P(f)$, that maximizes (4.35) is found via the Lagrangian technique. The resulting optimal power allocation is water-filling over frequency:

$$\frac{P(f)}{P} = \begin{cases} 1/\gamma_0 - 1/\gamma(f) & \gamma(f) \geq \gamma_0, \\ 0 & \gamma(f) < \gamma_0, \end{cases} \tag{4.36}$$

[3] We define γ_j with respect to the total power P because of the normalization by P on the left hand side of (4.32).

where $\gamma(f) = |H(f)|^2 P/N_0$. This results in channel capacity

$$C = \int_{f:\gamma(f)\geq\gamma_0} \log_2\left(\frac{\gamma(f)}{\gamma_0}\right)\, df. \tag{4.37}$$

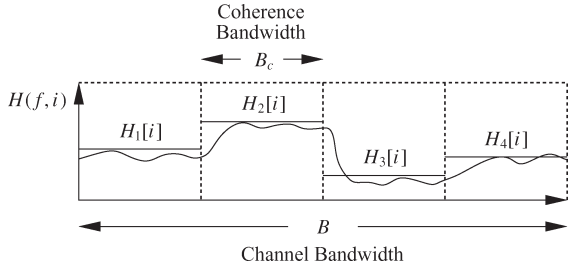


Figure 4.12: Channel division in frequency-selective fading.

Example 4.7: Consider a time-invariant frequency-selective block fading channel that has three subchannels of bandwidth $B = 1$ MHz. The frequency responses associated with each subchannel are $H_1 = 1,\ H_2 = 2$, and $H_3 = 3$, respectively. The transmit power constraint is $P = 10$ mW and the noise PSD $N_0/2$ has $N_0 = 10^{-9}$ W/Hz. Find the Shannon capacity of this channel and the optimal power allocation that achieves this capacity.
Solution: We first find $\gamma_j = |H_j|^2 P/(N_0 B)$ for each subchannel, yielding $\gamma_1 = 10,\ \gamma_2 = 40$, and $\gamma_3 = 90$. The cutoff γ_0 must satisfy (4.33). Assuming that all subchannels are allocated power, this yields

$$\frac{3}{\gamma_0} = 1 + \sum_j \frac{1}{\gamma_j} = 1.14 \Rightarrow \gamma_0 = 2.64 < \gamma_j,\ \forall j. \tag{4.38}$$

Since the cutoff γ_0 is less than γ_j for all j, our assumption that all subchannels are allocated power is consistent, so this is the correct cutoff value. The corresponding capacity is $C = \sum_{j=1}^{3} B\log_2(\gamma_j/\gamma_0) = 1000000(\log_2(10/2.64) + \log_2(40/2.64) + \log_2(90/2.64)) = 10.93$ Mbps.

4.3.2 Time-Varying Channels

The time-varying frequency-selective fading channel is similar to the model shown in Figure 4.9 except that $H(f) = H(f,\ i)$; that is, the channel varies over both frequency and time. It is difficult to determine the capacity of time-varying frequency-selective fading channels – even when the instantaneous channel $H(f,\ i)$ is known perfectly at the transmitter and receiver – because of the effects of intersymbol interference (ISI). In the case of transmitter and receiver side information, the optimal adaptation scheme must consider (a) the effect of the channel on the past sequence of transmitted bits and (b) how the ISI resulting from these bits will affect future transmissions [34]. The capacity of time-varying frequency-selective fading channels is in general unknown, but there do exist upper and lower bounds as well as limiting formulas [34, 35].

We can approximate channel capacity in time-varying frequency-selective fading by taking the channel bandwidth B of interest and then dividing it up into subchannels the size of the channel coherence bandwidth B_c, as shown in Figure 4.12. We then assume that each of the resulting subchannels is independent, time varying, and flat fading with $H(f,\ i) = H_j[i]$ on the jth subchannel.

Under this assumption, we obtain the capacity for each of these flat fading subchannels based on the average power $\bar{P}_j$ that we allocate to each subchannel, subject to a total power constraint $\bar{P}$. Since the channels are

independent, the total channel capacity is just equal to the sum of capacities on the individual narrowband flat fading channels – subject to the total average power constraint and averaged over both time and frequency:

$$C = \max_{\{\bar{P}_j\}:\sum_j \bar{P}_j \le \bar{P}} \sum_j C_j(\bar{P}_j), \tag{4.39}$$

where $C_j(\bar{P}_j)$ is the capacity of the flat fading subchannel with average power $\bar{P}_j$ and bandwidth B_c given by (4.17), (4.7), (4.25), or (4.30) for Shannon capacity under different side information and power allocation policies. We can also define $C_j(\bar{P}_j)$ as a capacity versus outage if only the receiver has side information.

We will focus on Shannon capacity assuming perfect transmitter and receiver channel CSI, since this upperbounds capacity under any other side information assumptions or suboptimal power allocation strategies. We know that if we fix the average power per subchannel then the optimal power adaptation follows a water-filling formula. We expect that the optimal average power to be allocated to each subchannel will also follow a water-filling formula, where more average power is allocated to better subchannels. Thus we expect that the optimal power allocation is a two-dimensional water-filling in both time and frequency. We now obtain this optimal two-dimensional water-filling and the corresponding Shannon capacity.

Define $\gamma_j[i] = |H_j[i]|^2\bar{P}/(N_0B)$ to be the instantaneous SNR on the jth subchannel at time i assuming the total power $\bar{P}$ is allocated to that time and frequency. We allow the power $P_j(\gamma_j)$ to vary with $\gamma_j[i]$. The Shannon capacity with perfect transmitter and receiver CSI is given by optimizing power adaptation relative to both time (represented by $\gamma_j[i] = \gamma_j$) and frequency (represented by the subchannel index j):

$$C = \max_{P_j(\gamma_j):\sum_j \int_0^\infty P_j(\gamma_j)p(\gamma_j)d\gamma_j \le \bar{P}} \sum_j \int_0^\infty B_c \log_2\left(1 + \frac{P_j(\gamma_j)\gamma_j}{\bar{P}}\right) p(\gamma_j)d\gamma_j. \tag{4.40}$$

To find the optimal power allocation $P_j(\gamma_j)$, we form the Lagrangian

$$J(P_j(\gamma_j)) = \sum_j \int_0^\infty B_c \log_2\left(1 + \frac{P_j(\gamma_j)\gamma_j}{\bar{P}}\right) p(\gamma_j)d\gamma_j - \lambda \sum_j \int_0^\infty P_j(\gamma_j)p(\gamma_j)d\gamma_j. \tag{4.41}$$

Note that (4.41) is similar to the Lagrangian (4.14) for the flat fading channel except that the dimension of frequency has been added by summing over the subchannels. Differentiating the Lagrangian and setting this derivative equal to zero eliminates all terms except the given subchannel and associated SNR:

$$\frac{\partial J(P_j(\gamma_j))}{\partial P_j(\gamma_j)} = \left[\left(\frac{B_c/\ln(2)}{1 + \gamma_j P_j(\gamma_j)/\bar{P}}\right)\frac{\gamma_j}{\bar{P}} - \lambda\right] p(\gamma_j) = 0. \tag{4.42}$$

Solving for $P_j(\gamma_j)$ yields the same water-filling as the flat-fading case:

$$\frac{P_j(\gamma_j)}{\bar{P}} = \begin{cases} 1/\gamma_0 - 1/\gamma_j & \gamma_j \ge \gamma_0, \\ 0 & \gamma_j < \gamma_0, \end{cases} \tag{4.43}$$

where the cutoff value γ_0 is obtained from the total power constraint over both time and frequency:

$$\sum_j \int_0^\infty P_j(\gamma_j)p(\gamma_j)d\gamma_j = \bar{P}. \tag{4.44}$$

Thus, the optimal power allocation (4.43) is a two-dimensional water-filling with a common cutoff value γ_0. Dividing the constraint (4.44) by $\bar{P}$ and substituting into the optimal power allocation (4.43), we get that γ_0 must satisfy

$$\sum_j \int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_j} \right) p(\gamma_j) d\gamma_j = 1. \tag{4.45}$$

It is interesting to note that, in the two-dimensional water-filling, the cutoff value for all subchannels is the same. This implies that even if the fading distribution or average fade power on the subchannels is different, all subchannels suspend transmission when the instantaneous SNR falls below the common cutoff value γ_0. Substituting the optimal power allocation (4.44) into the capacity expression (4.40) yields

$$C = \sum_j \int_{\gamma_0}^{\infty} B_c \log_2 \left(\frac{\gamma_j}{\gamma_0} \right) p(\gamma_j) d\gamma_j. \tag{4.46}$$

Chapter 4 Problems

1. Capacity in AWGN is given by $C = B\log_2(1 + P/(N_0B))$, where P is the received signal power, B is the signal bandwidth, and $N_0/2$ is the noise PSD. Find capacity in the limit of infinite bandwidth $B \to \infty$ as a function of P.

2. Consider an AWGN channel with bandwidth 50 MHz, received signal power 10 mW, and noise PSD $N_0/2$ where $N_0 = 2 \cdot 10^{-9}$ W/Hz. How much does capacity increase by doubling the received power? How much does capacity increase by doubling the channel bandwidth?

3. Consider two users simultaneously transmitting to a single receiver in an AWGN channel. This is a typical scenario in a cellular system with multiple users sending signals to a base station. Assume the users have equal received power of 10 mW and total noise at the receiver in the bandwidth of interest of 0.1 mW. The channel bandwidth for each user is 20 MHz.

   (a) Suppose that the receiver decodes user 1's signal first. In this decoding, user 2's signal acts as noise (assume it has the same statistics as AWGN). What is the capacity of user 1's channel with this additional interference noise?

   (b) Suppose that, after decoding user 1's signal, the decoder re-encodes it and subtracts it out of the received signal. Now, in the decoding of user 2's signal, there is no interference from user 1's signal. What then is the Shannon capacity of user 2's channel?

   Note: We will see in Chapter 14 that the decoding strategy of successively subtracting out decoded signals is optimal for achieving Shannon capacity of a multiuser channel with independent transmitters sending to one receiver.

4. Consider a flat fading channel of bandwidth 20 MHz and where, for a fixed transmit power $\bar{P}$, the received SNR is one of six values: $\gamma_1 = 20\,\text{dB},\ \gamma_2 = 15\,\text{dB},\ \gamma_3 = 10\,\text{dB},\ \gamma_4 = 5\,\text{dB},\ \gamma_5 = 0\,\text{dB}$, and $\gamma_6 = -5\,\text{dB}$. The probabilities associated with each state are $p_1 = p_6 = .1,\ p_2 = p_4 = .15$, and $p_3 = p_5 = .25$. Assume that only the receiver has CSI.

   (a) Find the Shannon capacity of this channel.

   (b) Plot the capacity versus outage for $0 \le P_{\text{out}} < 1$ and find the maximum average rate that can be correctly received (maximum C_{out}).

5. Consider a flat fading channel in which, for a fixed transmit power $\bar{P}$, the received SNR is one of four values: $\gamma_1 = 30\,\text{dB},\ \gamma_2 = 20\,\text{dB},\ \gamma_3 = 10\,\text{dB}$, and $\gamma_4 = 0\,\text{dB}$. The probabilities associated with each state are $p_1 = .2,\ p_2 = .3,\ p_3 = .3$, and $p_4 = .2$. Assume that both transmitter and receiver have CSI.

   (a) Find the optimal power adaptation policy $P[i]/\bar{P}$ for this channel and its corresponding Shannon capacity per unit h Hertz (C/B).

   (b) Find the channel inversion power adaptation policy for this channel and associated zero-outage capacity per unit bandwidth.

   (c) Find the truncated channel inversion power adaptation policy for this channel and associated outage capacity per unit bandwidth for three different outage probabilities: $P_{\text{out}} = .1,\ P_{\text{out}} = .25$, and P_{out} (and the associated cutoff γ_0) equal to the value that achieves maximum outage capacity.

6. Consider a cellular system where the power falloff with distance follows the formula $P_r(d) = P_t(d_0/d)^\alpha$, where $d_0 = 100\,\mathrm{m}$ and α is a random variable. The distribution for α is $p(\alpha = 2) = .4,\ p(\alpha = 2.5) = .3,\ p(\alpha = 3) = .2$, and $p(\alpha = 4) = .1$. Assume a receiver at a distance $d = 1000\,\mathrm{m}$ from the transmitter, with an average transmit power constraint of $P_t = 100\,\mathrm{mW}$ and a receiver noise power of .1 mW. Assume that both transmitter and receiver have CSI.

   (a) Compute the distribution of the received SNR.

   (b) Derive the optimal power adaptation policy for this channel and its corresponding Shannon capacity per unit Hertz (C/B).

   (c) Determine the zero-outage capacity per unit bandwidth of this channel.

   (d) Determine the maximum outage capacity per unit bandwidth of this channel.

7. In this problem we explore power allocation under a fixed energy constraint to maximize the number of bits transmitted. Consider a cellular system with a mobile receiver. The receiver stops at point A (100m from the cell phone tower) for 10 minutes and point B (500m from the cell phone tower) for 20 minutes. Assume the signal from the cell phone tower to the receiver only experiences path loss under the simplified path loss model with $\gamma = 3,\ d_0 = 1$m, and $K = 1$ along with AWGN, that the channel bandwidth is 20MHz, and the receiver noise PSD is $N_0 = 5 \times 10^{-14}$mW/Hz. Also, assume a fixed transmit power of 1mW.

   (a) What is the SNR received at points A and B?

   (b) What is total amount of energy (in Joules) that is transmitted during the two stops.

   (c) Assuming the transmitter sends data with a rate equal to the Shannon capacity of the channel at each of the stops, what is the total number of bits received?

   (d) Suppose now that the cell phone tower does not send at a fixed transmit power but fixes the total number of Joules it will transmit to the receiver at each of the two stops. Using the total amount of energy derived in part (b), determine the optimal allocation of this total energy at each of the stops to maximize the number of bits received. Assume that the transmitter sends data at a rate equal to the Shannon capacity with SNR based on your optimal energy allocation. How many bits are received with this optimal energy allocation and how does it compare to your answer in part (c)?

8. In this problem we explore why channel inversion leads to zero capacity in Rayleigh fading. Assume a Rayleigh fading channel with $\overline{\gamma} = 1$ so the receiver power has an exponential distribution $p(\gamma) = e^{-\gamma}$.

   (a) Show using a first-order Taylor series expansion of $e^{-\gamma}$ around $\gamma = 0$ that $p(\gamma) = e^{-\gamma} > 1 - \gamma$.

   (b) We will now show that $\mathbf{E}\left[\frac{1}{\gamma}\right]$ is unbounded by taking the expectation for $\gamma \in (0, c)$ for some small constant c. Fix this small constant c and a smaller constant $\epsilon < c$. Then compute the following lower bound for $\mathbf{E}\left[\frac{1}{\gamma}\right]$ by taking the expectation over the interval (ϵ, c) and using the lower bound on $p(\gamma)$ from part (a):
   $$\mathbf{E}\left[\frac{1}{\gamma}\right] > \int_\epsilon^c \frac{1}{\gamma}(1-\gamma)d\gamma.$$
   Show that as $\epsilon \to 0$, the above expression becomes unbounded. This suggests that the transmit power needed to maintain a constant received power by inverting Rayleigh fading is unbounded.

9. Assume a Rayleigh fading channel, where the transmitter and receiver have CSI and the distribution of the fading SNR $p(\gamma)$ is exponential with mean $\bar{\gamma} = 10$ dB. Assume a channel bandwidth of 10 MHz.

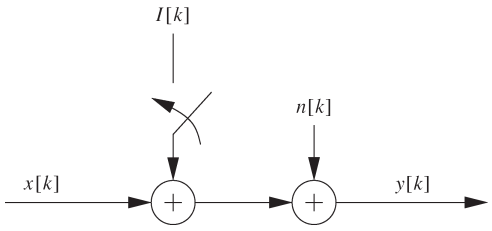


Figure 4.13: Interference channel for Problem 4.8.

(a) Find the cutoff value γ_0 and the corresponding power adaptation that achieves Shannon capacity on this channel.

(b) Compute the Shannon capacity of this channel.

(c) Compare your answer in part (b) with the channel capacity in AWGN with the same average SNR.

(d) Compare your answer in part (b) with the Shannon capacity when only the receiver knows $\gamma[i]$.

(e) Compare your answer in part (b) with the zero-outage capacity and outage capacity when the outage probability is .05.

(f) Repeat parts (b), (c), and (d) – that is, obtain the Shannon capacity with perfect transmitter and receiver side information, in AWGN for the same average power, and with just receiver side information – for the same fading distribution but with mean $\bar{\gamma} = -5$ dB. Describe the circumstances under which a fading channel has higher capacity than an AWGN channel with the same average SNR and explain why this behavior occurs.

10. This problem illustrates the capacity gains that can be obtained from interference estimation and also how a malicious jammer can wreak havoc on link performance. Consider the interference channel depicted in Figure 4.13. The channel has a combination of AWGN $n[k]$ and interference $I[k]$. We model $I[k]$ as AWGN. The interferer is on (i.e., the switch is down) with probability .25 and off (i.e., switch up) with probability .75. The average transmit power is 10 mW, the noise PSD has $N_0 = 10^{-8}$ W/Hz, the channel bandwidth B is 10 kHz (receiver noise power is N_0B), and the interference power (when on) is 9 mW.

(a) What is the Shannon capacity of the channel if neither transmitter nor receiver know when the interferer is on?

(b) What is the capacity of the channel if both transmitter and receiver know when the interferer is on?

(c) Suppose now that the interferer is a malicious jammer with perfect knowledge of $x[k]$ (so the interferer is no longer modeled as AWGN). Assume that neither transmitter nor receiver has knowledge of the jammer behavior. Assume also that the jammer is always on and has an average transmit power of 10 mW. What strategy should the jammer use to minimize the SNR of the received signal?

11. Consider the malicious interferer of Problem 4.8. Suppose that the transmitter knows the interference signal perfectly. Consider two possible transmit strategies under this scenario: the transmitter can ignore the interference and use all its power for sending its signal, or it can use some of its power to cancel out the interferer (i.e., transmit the negative of the interference signal). In the first approach the interferer will degrade capacity by increasing the noise, and in the second strategy the interferer also degrades capacity because the transmitter sacrifices some power to cancel out the interference. Which strategy results in higher capacity? *Note*: There is a third strategy, in which the encoder actually exploits the structure of the interference in its encoding. This strategy is called *dirty paper coding* and is used to achieve Shannon capacity on broadcast channels with multiple antennas, as described in Chapter 14.

12. Show using Lagrangian techniques that the optimal power allocation to maximize the capacity of a time-invariant block fading channel is given by the water-filling formula in (4.32).

13. Consider a time-invariant block fading channel with frequency response

$$H(f) = \begin{cases} 1 & f_c - 20 \text{ MHz} \leq f < f_c - 10 \text{ MHz}, \\ .5 & f_c - 10 \text{ MHz} \leq f < f_c, \\ 2 & f_c \leq f < f_c + 10 \text{ MHz}, \\ .25 & f_c + 10 \text{ MHz} \leq f < f_c + 20 \text{ MHz}, \\ 0 & \text{else}, \end{cases}$$

for $f > 0$ and $H(-f) = H(f)$. For a transmit power of 10 mW and a noise PSD of $.001\,\mu$W per Hertz, find the optimal power allocation and corresponding Shannon capacity of this channel.

14. Show that the optimal power allocation to maximize the capacity of a time-invariant frequency-selective fading channel is given by the water-filling formula in (4.36).

15. Consider a frequency-selective fading channel with total bandwidth 12 MHz and coherence bandwidth $B_c = 4$ MHz. Divide the total bandwidth into three subchannels of bandwidth B_c, and assume that each subchannel is a Rayleigh flat fading channel with independent fading on each subchannel. Assume the subchannels have average gains $\mathbf{E}[|H_1[i]|^2] = 1,\ \mathbf{E}[|H_2[i]|^2] = .5$, and $\mathbf{E}[|H_3[i]|^2] = .125$. Assume a total transmit power of 30 mW and a receiver noise PSD with $N_0 = .001\,\mu$W/Hz.

    (a) Find the optimal two-dimensional water-filling power adaptation for this channel and the corresponding Shannon capacity, assuming both transmitter and receiver know the instantaneous value of $H_j[i],\ j = 1, 2, 3$.

    (b) Compare the capacity derived in part (a) with that obtained by allocating an equal average power of 10 mW to each subchannel and then water-filling on each subchannel relative to this power allocation.

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, pp. 379–423, 623–56, 1948.

[2] C. E. Shannon, "Communications in the presence of noise." *Proc. IRE*, pp. 10–21, 1949.

[3] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

[4] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, pp. 933–46, May 2000.

[5] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[6] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 2nd Ed, 2006.

[7] Y. Polyanskiy, H. V. Poor and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307-2359, May 2010.

[8] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes", *IEEE Transactions on Commununications*, vol. 44, no. 10, pp. 1261-1271, Oct. 1996.

[9] R.G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963 (Sc.D. MIT, 1960).

[10] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel", *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 273-278, Jan. 1998.

[11] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels", *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051-3073, July 2009.

[12] I. Csiszár and J. Körner, *Information Theory*: *Coding Theorems for Discrete Memoryless Channels*, Academic Press, New York, 1981.

[13] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel," *IEEE Trans. Inform. Theory*, pp. 18–26, January 1991.

[14] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh fading channels," *IEEE Trans. Inform. Theory*, pp. 1290–1301, May 2001.

[15] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inform. Theory*, pp. 2426–67, October 2003.

[16] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inform. Theory*, pp. 868–86, May 1996.

[17] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliot channel," *IEEE Trans. Inform. Theory*, pp. 1277–90, November 1989.

[18] T. Holliday, A. Goldsmith, and P. Glynn, "Capacity of finite state Markov channels with general inputs," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 289, July 2003.

[19] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, pp. 44–53, January 1984.

[20] A. Lapidoth and S. Shamai, "Fading channels: How perfect need 'perfect side information' be?" *IEEE Trans. Inform. Theory*, pp. 1118–34, November 1997.

[21] G. J. Foschini, D. Chizhik, M. Gans, C. Papadias, and R. A. Valenzuela, "Analysis and performance of some basic space-time architectures," *IEEE J. Sel. Areas Commun.*, pp. 303–20, April 2003.

[22] W. L. Root and P. P. Varaiya, "Capacity of classes of Gaussian channels," *SIAM J. Appl. Math.*, pp. 1350–93, November 1968.

[23] J. Wolfowitz, *Coding Theorems of Information Theory*, 2nd ed., Springer-Verlag, New York, 1964.

[24] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, pp. 1986–92, November 1997.

[25] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, pp. 2007–19, September 1999.

[26] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity combining techniques," *IEEE Trans. Veh. Tech.* pp. 1165–81, July 1999.

[27] S.-G. Chua and A. J. Goldsmith, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, pp. 1218–30, October 1997.

[28] S.-G. Chua and A. J. Goldsmith, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, pp. 595–602, May 1998.

[29] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech.* pp. 303–12, May 1991.

[30] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multi-element antennas," *Bell System Tech. J.*, pp. 41–59, Autumn 1996.

[31] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Labs Internal Tech. Memo, June 1995.

[32] G. J. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, pp. 311–35, March 1998.

[33] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inform. Theory*, pp. 380–8, May 1988.

[34] A. Goldsmith and M. Medard, "Capacity of time-varying channels with channel side information," *IEEE Trans. Inform. Theory* Vol. 53, No. 3, pp. 881-899, March 2007.

[35] S. Diggavi, "Analysis of multicarrier transmission in time-varying channels," *Proc. IEEE Internat. Conf. Commun*., pp. 1191–5, June 1997.

Chapter 5

Digital Modulation and Detection

The advances over the last several decades in hardware and digital signal processing have made digital transceivers much cheaper, faster, and more power efficient than analog transceivers. More importantly, digital modulation offers a number of other advantages over analog modulation, including higher spectral efficiency, powerful error correction techniques, resistance to channel impairments, more efficient multiple access strategies, and better security and privacy. Specifically, high-level digital modulation techniques such as MQAM allow much more efficient use of spectrum than is possible with analog modulation. Advances in coding and coded modulation applied to digital signaling make the signal much less susceptible to noise and fading, and equalization or multicarrier techniques can be used to mitigate intersymbol interference (ISI). Spread-spectrum techniques applied to digital modulation can simultaneously remove or combine multipath, resist interference, and detect multiple users. Finally, digital modulation is much easier to encrypt, resulting in a higher level of security and privacy for digital systems. For all these reasons, systems currently being built or proposed for wireless applications are all digital systems.

Digital modulation and detection consist of transferring information in the form of bits over a communication channel. The bits are binary digits taking on the values of either 1 or 0. These information bits are derived from the information source, which may be a digital source or an analog source that has been passed through an A/D converter. Both digital and A/D-converted analog sources may be compressed to obtain the information bit sequence. Digital modulation consists of mapping the information bits into an analog signal for transmission over the channel. Detection consists of estimating the original bit sequence based on the signal received over the channel. The main considerations in choosing a particular digital modulation technique are high data rate; high spectral efficiency (minimum bandwidth occupancy); high power efficiency (minimum required transmit power); robustness to channel impairments (minimum probability of bit error); and low power/cost implementation. Often these are conflicting requirements, and the choice of modulation is based on finding the technique that achieves the best trade-off between them.

There are two main categories of digital modulation: amplitude/phase modulation and frequency modulation. Since frequency modulation typically has a constant signal envelope and is generated using nonlinear techniques, this modulation is also called *constant envelope modulation* or *nonlinear modulation*, and amplitude/phase modulation is also called *linear modulation*. Linear modulation generally has better spectral properties than nonlinear modulation, since nonlinear processing leads to spectral broadening. However, amplitude and phase modulation embeds the information bits into the amplitude or phase of the transmitted signal, which is more susceptible to variations from fading and interference. In addition, amplitude and phase modulation techniques typically require linear amplifiers, which are more expensive and less power efficient than the nonlinear amplifiers that can be used with nonlinear modulation. Thus, the trade-off between linear versus nonlinear modulation is one of better spectral efficiency for the former technique and better power efficiency and resistance to channel impairments for the latter. Once the modulation technique is determined, the constellation size must be chosen. Modulations with large

constellations have higher data rates for a given signal bandwidth, but they are more susceptible to noise, fading, and hardware imperfections. Finally, some demodulators require a coherent phase reference with respect to the transmitted signal. Obtaining this coherent reference may be difficult or significantly increase receiver complexity. Thus, modulation techniques that do not require a coherent phase reference in the receiver are desirable.

We begin this chapter with a general discussion of signal space concepts. These concepts greatly simplify the design and analysis of modulation and demodulation techniques by mapping infinite-dimensional signals to a finite-dimensional vector space. The general principles of signal space analysis will then be applied to the analysis of amplitude and phase modulation techniques, including pulse amplitude modulation (PAM), phase-shift keying (PSK), and quadrature amplitude modulation (QAM). We will also discuss constellation shaping and quadrature offset techniques for these modulations, as well as differential encoding to avoid the need for a coherent phase reference. We then describe frequency modulation techniques and their properties, including frequency-shift keying (FSK), minimum-shift keying (MSK), and continuous-phase FSK (CPFSK). Both coherent and noncoherent detection of these techniques will be discussed. Pulse-shaping techniques to improve the spectral properties of the modulated signals will also be covered, along with issues associated with carrier phase recovery and symbol synchronization.

5.1 Signal Space Analysis

Digital modulation encodes a bit stream of finite length into one of several possible transmitted signals. Intuitively, the receiver minimizes the probability of detection error by decoding the received signal as the signal in the set of possible transmitted signals that is "closest" to the one received. Determining the distance between the transmitted and received signals requires a metric for the distance between signals. By representing signals as projections onto a set of basis functions, we obtain a one-to-one correspondence between the set of transmitted signals and their vector representations. Thus, we can analyze signals in finite-dimensional vector space instead of infinite-dimensional function space, using classical notions of distance for vector spaces. In this section we show how digitally modulated signals can be represented as vectors in an appropriately defined vector space and how optimal demodulation methods can be obtained from this vector-space representation. This general analysis will then be applied to specific modulation techniques in later sections.

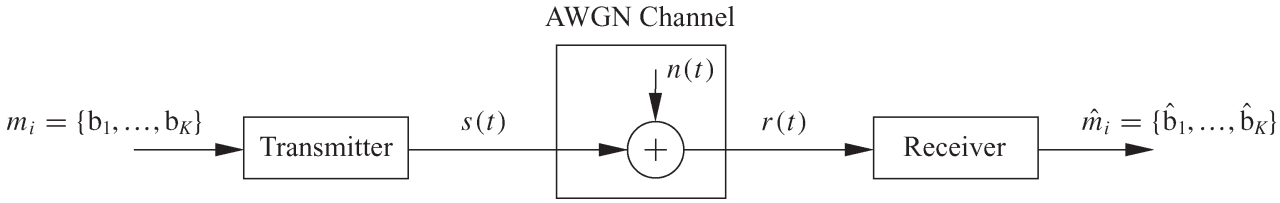


Figure 5.1: Communication system model.

5.1.1 Signal and System Model

Consider the communication system model shown in Figure 5.1. Every T seconds, the system sends $K = \log_2 M$ bits of information through the channel for a data rate of $R = K/T$ bits per second (bps). There are $M = 2^K$ possible sequences of K bits, and we say that each bit sequence of length K comprises a message $m_i = \{b_1, \ldots, b_K\} \in \mathcal{M}$, where $\mathcal{M} = \{m_1, \ldots, m_M\}$ is the set of all such messages. The messages have probability p_i of being selected for transmission, where $\sum_{i=1}^{M} p_i = 1$.

Suppose message m_i is to be transmitted over the channel during the time interval $[0, T)$. Because the channel is analog, the message must be embedded into an analog signal for channel transmission. Hence, each message

$m_i \in \mathcal{M}$ is mapped to a unique analog signal $s_i(t) \in S = \{s_1(t), \ldots, s_M(t)\}$, where $s_i(t)$ is defined on the time interval $[0, T)$ and has energy

$$E_{s_i} = \int_0^T s_i^2(t)dt, \quad i = 1, \ldots, M. \tag{5.1}$$

Since each message represents a bit sequence it follows that each signal $s_i(t) \in S$ also represents a bit sequence, and detection of the transmitted signal $s_i(t)$ at the receiver is equivalent to detection of the transmitted bit sequence. When messages are sent sequentially, the transmitted signal becomes a sequence of the corresponding analog signals over each time interval $[kT, (k+1)T) : s(t) = \sum_k s_i(t-kT)$, where $s_i(t)$ is a baseband or passband analog signal corresponding to the message m_i designated for the transmission interval $[kT, (k+1)T)$. This is illustrated in Figure 5.2, where we show the transmitted signal $s(t) = s_1(t) + s_2(t-T) + s_1(t-2T) + s_1(t-3T)$ corresponding to the string of messages m_1, m_2, m_1, m_1 with message m_i mapped to signal $s_i(t)$.

In the model of Figure 5.1, the transmitted signal is sent through an AWGN channel, where a white Gaussian noise process $n(t)$ of power spectral density $N_0/2$ is added to form the received signal $r(t) = s(t) + n(t)$. Given $r(t)$, the receiver must determine the best estimate of which $s_i(t) \in S$ was transmitted during each transmission interval $[kT, (k+1)T)$. This best estimate for $s_i(t)$ is mapped to a best estimate of the message $m_i \in \mathcal{M}$ and the receiver then outputs this best estimate $\hat{m} = \{\hat{\mathrm{b}}_1, \ldots, \hat{\mathrm{b}}_K\} \in \mathcal{M}$ of the transmitted bit sequence.

The goal of the receiver design in estimating the transmitted message is to minimize the probability of message error,

$$P_e = \sum_{i=1}^{M} p(\hat{m} \neq m_i | m_i \text{ sent}) p(m_i \text{ sent}), \tag{5.2}$$

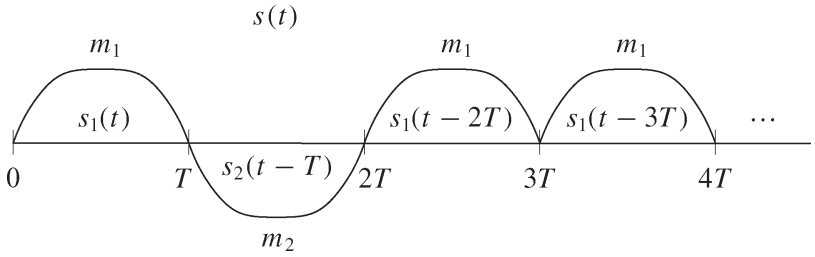


Figure 5.2: Transmitted signal for a sequence of messages.

over each time interval $[kT, (k+1)T)$. By representing the signals $\{s_i(t) : i = 1, \ldots, M\}$ geometrically, we can solve for the optimal receiver design in AWGN based on a minimum distance criterion. Note that, as we saw in previous chapters, wireless channels typically have a time-varying impulse response in addition to AWGN. We will consider the effect of an arbitrary channel impulse response on digital modulation performance in Chapter 6; methods to combat this performance degradation are discussed in Chapters 11–13.

5.1.2 Geometric Representation of Signals

The basic premise behind a geometrical representation of signals is the notion of a basis set. Specifically, using a Gram–Schmidt orthogonalization procedure [1, 2], it can be shown that any set of M real signals $S = \{s_1(t), \ldots, s_M(t)\}$ defined on $[0, T)$ with finite energy can be represented as a linear combination of $N \leq M$ real orthonormal basis functions $\{\phi_1(t), \ldots, \phi_N(t)\}$. We say that these basis functions *span* the set S. Thus, we can write each $s_i(t) \in S$ in terms of its *basis function representation* as

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t), \quad 0 \le t < T, \tag{5.3}$$

where

$$s_{ij} = \int_0^T s_i(t)\phi_j(t)dt \tag{5.4}$$

is a real coefficient representing the projection of $s_i(t)$ onto the basis function $\phi_j(t)$ and

$$\int_0^T \phi_i(t)\phi_j(t)dt = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{5.5}$$

If the signals $\{s_i(t)\}$ are linearly independent then $N = M$, otherwise $N < M$. Moreover, the minimum number N of orthogonal basis functions needed to represent any signal $s_i(t)$ of duration T and bandwidth B is roughly $2BT$ [3, Chap. 5.3]. The signal $s_i(t)$ thus occupies $2BT$ orthogonal dimensions.

For linear passband modulation techniques, the basis set consists of the sine and cosine functions:

$$\phi_1(t) = \sqrt{\frac{2}{T}}\cos(2\pi f_c t) \tag{5.6}$$

and

$$\phi_2(t) = \sqrt{\frac{2}{T}}\sin(2\pi f_c t). \tag{5.7}$$

The $\sqrt{2/T}$ factor is needed for normalization so that $\int_0^T \phi_i^2(t)\,dt = 1, i = 1, 2$. In fact, with these basis functions we get only an approximation to (5.5), since

$$\int_0^T \phi_1^2(t)\,dt = \frac{2}{T}\int_0^T .5[1 + \cos(4\pi f_c t)]\,dt = 1 + \frac{\sin(4\pi f_c T)}{4\pi f_c T}. \tag{5.8}$$

The numerator in the second term of (5.8) is bounded by 1 and for $f_c T \gg 1$ the denominator of this term is very large. Hence this second term can be neglected. Similarly,

$$\int_0^T \phi_1(t)\phi_2(t)\,dt = \frac{2}{T}\int_0^T .5\sin(4\pi f_c t)\,dt = \frac{1 - \cos(4\pi f_c T)}{4\pi f_c T} \approx 0, \tag{5.9}$$

where the approximation is taken as an equality for $f_c T \gg 1$.

With the basis set $\phi_1(t) = \sqrt{2/T}\cos(2\pi f_c t)$ and $\phi_2(t) = \sqrt{2/T}\sin(2\pi f_c t)$, the basis function representation (5.3) corresponds to the equivalent lowpass representation of $s_i(t)$ in terms of its in-phase and quadrature components:

$$s_i(t) = s_{i1}\sqrt{\frac{2}{T}}\cos(2\pi f_c t) + s_{i2}\sqrt{\frac{2}{T}}\sin(2\pi f_c t). \tag{5.10}$$

Note that the carrier basis functions may have an initial phase offset ϕ_0. The basis set may also include a baseband pulse-shaping filter $g(t)$ to improve the spectral characteristics of the transmitted signal:

$$s_i(t) = s_{i1}g(t)\cos(2\pi f_c t) + s_{i2}g(t)\sin(2\pi f_c t). \tag{5.11}$$

In this case the pulse shape $g(t)$ must maintain the orthonormal properties (5.5) of basis functions; that is, we must have

$$\int_0^T g^2(t)\cos^2(2\pi f_c t)\,dt = 1 \tag{5.12}$$

and

$$\int_0^T g^2(t)\cos(2\pi f_c t)\sin(2\pi f_c t)\,dt = 0, \tag{5.13}$$

where the equalities may be approximations for $f_c T \gg 1$ as in (5.8) and (5.9). If the bandwidth of $g(t)$ satisfies $B \ll f_c$ then $g^2(t)$ is roughly constant over $T_c = 1/f_c$, so (5.13) is approximately true because the sine and cosine functions are orthogonal over one period T_c. The simplest pulse shape that satisfies (5.12) and (5.13) is the rectangular pulse shape $g(t) = \sqrt{2/T}, 0 \le t < T$.

Example 5.1: Binary phase-shift keying (BPSK) modulation transmits the signal $s_1(t) = \alpha\cos(2\pi f_c t)$, $0 \le t \le T$, to send a 1-bit and the signal $s_2(t) = -\alpha\cos(2\pi f_c t)$, $0 \le t \le T$, to send a 0-bit. Find the set of orthonormal basis functions and coefficients $\{s_{ij}\}$ for this modulation.
Solution: There is only one basis function for $s_1(t)$ and $s_2(t)$:

$$\phi(t) = \sqrt{2/T}\cos(2\pi f_c t), \tag{5.14}$$

where the $\sqrt{2/T}$ is needed for normalization. The coefficients are then given by $s_1 = \alpha\sqrt{T/2}$ and $s_2 = -\alpha\sqrt{T/2}$.

Let $\mathbf{s}_i = (s_{i1}, \ldots, s_{iN}) \in \mathbb{R}^N$ be the vector of coefficients $\{s_{ij}\}$ in the basis representation of $s_i(t)$. We call $\mathbf{s}_i$ the *signal constellation point* corresponding to the signal $s_i(t)$. The *signal constellation* consists of all constellation points $\{\mathbf{s}_1, \ldots, \mathbf{s}_M\}$. Given the basis functions $\{\phi_1(t), \ldots, \phi_N(t)\}$, there is a one-to-one correspondence between the transmitted signal $s_i(t)$ and its constellation point $\mathbf{s}_i$. Specifically, $s_i(t)$ can be obtained from $\mathbf{s}_i$ by (5.3) and $\mathbf{s}_i$ can be obtained from $s_i(t)$ by (5.4). Thus, it is equivalent to characterize the transmitted signal by $s_i(t)$ or $\mathbf{s}_i$. The representation of $s_i(t)$ in terms of its constellation point $\mathbf{s}_i \in \mathbb{R}^N$ is called its *signal space representation*, and the vector space containing the constellation is called the *signal space*. A two-dimensional signal space is illustrated in Figure 5.3, where we show $\mathbf{s}_i \in \mathbb{R}^2$ with the ith axis of $\mathbb{R}^2$ corresponding to the basis function $\phi_i(t)$, $i = 1, 2$. With this signal space representation we can analyze the infinite-dimensional functions $s_i(t)$ as vectors $\mathbf{s}_i$ in finite-dimensional vector space $\mathbb{R}^2$. This greatly simplifies the analysis of system performance as well as the derivation of optimal receiver design. Signal space representations for common modulation techniques like MPSK and MQAM are two-dimensional (corresponding to the in-phase and quadrature basis functions) and will be given later in the chapter.

In order to analyze signals via a signal space representation, we require a few definitions for vector characterization in the vector space $\mathbb{R}^N$. The length of a vector in $\mathbb{R}^N$ is defined as

$$\|\mathbf{s}_i\| \triangleq \sqrt{\sum_{j=1}^{N} s_{ij}^2}. \tag{5.15}$$

The distance between two signal constellation points $\mathbf{s}_i$ and $\mathbf{s}_k$ is thus

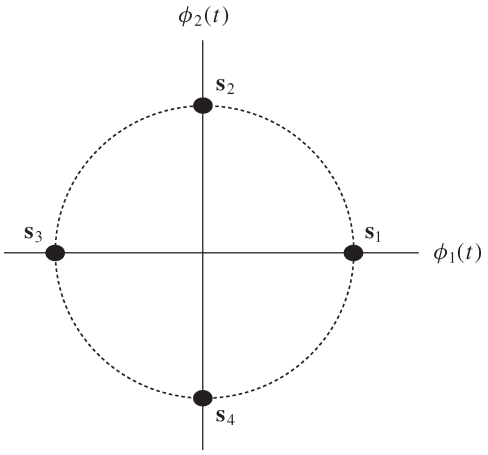


Figure 5.3: Signal space representation.

$$\|\mathbf{s}_i - \mathbf{s}_k\| = \sqrt{\sum_{j=1}^{N}(s_{ij} - s_{kj})^2} = \sqrt{\int_0^T (s_i(t) - s_k(t))^2 dt}, \tag{5.16}$$

where the second equality is obtained by writing $s_i(t)$ and $s_k(t)$ in their basis representation (5.3) and using the orthonormal properties of the basis functions. Finally, the inner product $\langle s_i(t),\, s_k(t)\rangle$ between two real signals $s_i(t)$ and $s_k(t)$ on the interval $[0, T]$ is

$$\langle s_i(t), s_k(t)\rangle = \int_0^T s_i(t)s_k(t)\, dt. \tag{5.17}$$

Similarly, the inner product $\langle \mathbf{s}_i, \mathbf{s}_k\rangle$ between two constellation points is

$$\langle \mathbf{s}_i, \mathbf{s}_k\rangle = \mathbf{s}_i\mathbf{s}_k^T = \int_0^T s_i(t)s_k(t)\, dt = \langle s_i(t), s_k(t)\rangle, \tag{5.18}$$

where the equality between the vector inner product and the corresponding signal inner product follows from the basis representation of the signals (5.3) and the orthonormal property of the basis functions (5.5). We say that two signals are *orthogonal* if their inner product is zero. Thus, by (5.5), the basis functions are orthogonal functions.

5.1.3 Receiver Structure and Sufficient Statistics

Given the channel output $r(t) = s_i(t) + n(t)$, $0 \le t < T$, we now investigate the receiver structure to determine which constellation point $\mathbf{s}_i$ (or, equivalently, which message m_i) was sent over the time interval $[0, T)$. A similar procedure is done for each time interval $[kT, (k + 1)T)$. We would like to convert the received signal $r(t)$ over each time interval into a vector, as this would allow us to work in finite-dimensional vector space when estimating the transmitted signal. However, this conversion should not be allowed to compromise the estimation accuracy. We now study a receiver that converts the received signal to a vector without compromising performance. Consider the receiver structure shown in Figure 5.4, where

$$s_{ij} = \int_0^T s_i(t)\phi_j(t)\, dt \tag{5.19}$$

and

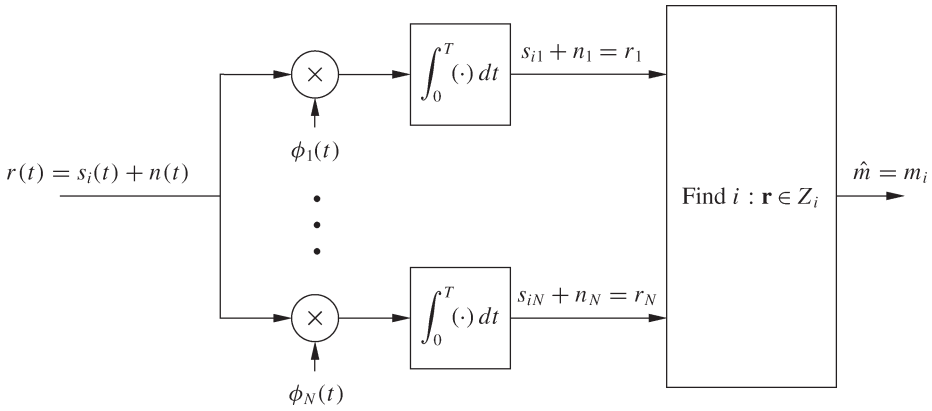


Figure 5.4: Receiver structure for signal detection in AWGN.

$$n_j = \int_0^T n(t)\phi_j(t)\, dt. \tag{5.20}$$

We can rewrite $r(t)$ as

$$\sum_{j=1}^{N}(s_{ij} + n_j)\phi_j(t) + n_r(t) = \sum_{j=1}^{N} r_j\phi_j(t) + n_r(t); \tag{5.21}$$

here $r_j = s_{ij} + n_j$ and $n_r(t) = n(t) - \sum_{j=1}^{N} n_j\phi_j(t)$ denotes the "remainder" noise, which is the component of the noise orthogonal to the signal space. If we can show that the optimal detection of the transmitted signal constellation point $\mathbf{s}_i$ given received signal $r(t)$ does not make use of the remainder noise $n_r(t)$, then the receiver can make its estimate $\hat{m}$ of the transmitted message m_i as a function of $\mathbf{r} = (r_1, \ldots, r_N)$ alone. In other words, $\mathbf{r} = (r_1, \ldots, r_N)$ is a *sufficient statistic* for $r(t)$ in the optimal detection of the transmitted messages. In that case, we only need to determine the decision region Z_i in which $\mathbf{r}$ falls, where the decision region will be defined later in Sec. 5.1.4.

It is intuitively clear that the remainder noise $n_r(t)$ should not help in detecting the transmitted signal $s_i(t)$, since its projection onto the signal space is zero. This is illustrated in Figure 5.5, where we assume the transmitted signal lies in a space spanned by the basis set $(\phi_1(t), \phi_2(t))$ while the remainder noise lies in a space spanned by the basis function $\phi_{n_r}(t)$, which is orthogonal to $\phi_1(t)$ and $\phi_2(t)$. Specifically, the remainder noise in Figure 5.5 is represented by n_r, where $n_r(t) = n_r\phi_{n_r}(t)$. The received signal is represented by $\mathbf{r} + n_r$, where $\mathbf{r} = (r_1, r_2)$ with $r(t) - n_r(t) = r_1\phi_1(t) + r_2\phi_2(t)$. From the figure it appears that projecting $\mathbf{r} + n_r$ onto $\mathbf{r}$ will not compromise the detection of which constellation $\mathbf{s}_i$ was transmitted, since n_r lies in a space orthogonal to the space in which $\mathbf{s}_i$ lies. We now proceed to show mathematically why this intuition is correct.

Let us first examine the distribution of $\mathbf{r}$. Since $n(t)$ is a Gaussian random process, if we condition on the transmitted signal $s_i(t)$ then the channel output $r(t) = s_i(t) + n(t)$ is also a Gaussian random process and $\mathbf{r} = (r_1, \ldots, r_N)$ is a Gaussian random vector. Recall that $r_j = s_{ij} + n_j$. Thus, conditioned on a transmitted constellation $\mathbf{s}_i$, we have that

$$\mu_{r_j|\mathbf{s}_i} = \mathbf{E}[r_j|\mathbf{s}_i] = \mathbf{E}[s_{ij} + n_j|s_{ij}] = s_{ij} \tag{5.22}$$

(since $n(t)$ has zero mean) and

$$\sigma_{r_j|\mathbf{s}_i} = \mathbf{E}[r_j - \mu_{r_j|\mathbf{s}_i}]^2 = \mathbf{E}[s_{ij} + n_j - s_{ij}|s_{ij}]^2 = \mathbf{E}[n_j^2]. \tag{5.23}$$

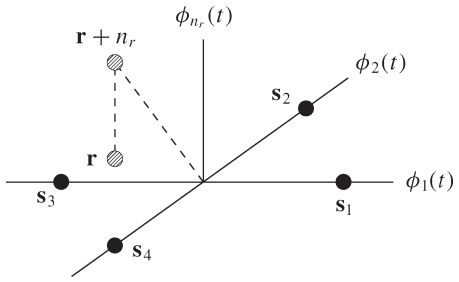


Figure 5.5: Projection of received signal onto received vector $\mathbf{r}$.

Moreover,

$$\begin{aligned}
\mathrm{Cov}[r_j r_k|\mathbf{s}_i] &= \mathbf{E}[(r_j-\mu_{r_j})(r_k-\mu_{r_k})|\mathbf{s}_i] \\
&= \mathbf{E}[n_j n_k] \\
&= \mathbf{E}\left[\int_0^T n(t)\phi_j(t)\,dt \int_0^T n(\tau)\phi_k(\tau)\,d\tau\right] \\
&= \int_0^T\int_0^T \mathbf{E}[n(t)n(\tau)]\phi_j(t)\phi_k(\tau)\,dt\,d\tau \\
&= \int_0^T\int_0^T \frac{N_0}{2}\delta(t-\tau)\phi_j(t)\phi_k(\tau)\,dt\,d\tau \\
&= \frac{N_0}{2}\int_0^T \phi_j(t)\phi_k(t)\,dt \\
&= \begin{cases} N_0/2 & j=k, \\ 0 & j\neq k, \end{cases}
\end{aligned} \tag{5.24}$$

where the last equality follows from the orthonormality of the basis functions. Thus, conditioned on the transmitted constellation $\mathbf{s}_i$, the r_j are uncorrelated and, since they are Gaussian, they are also independent. Furthermore, $\mathbf{E}[n_j^2] = N_0/2$.

We have shown that, conditioned on the transmitted constellation $\mathbf{s}_i$, r_j is a Gauss-distributed random variable that is independent of $r_k (k \neq j)$ with mean s_{ij} and variance $N_0/2$. Thus, the conditional distribution of $\mathbf{r}$ is given by

$$p(\mathbf{r}\,|\,\mathbf{s}_i \text{ sent}) = \prod_{j=1}^{N} p(r_j\,|\,m_i) = \frac{1}{(\pi N_0)^{N/2}} \exp\left[-\frac{1}{N_0}\sum_{j=1}^{N}(r_j - s_{ij})^2\right]. \tag{5.25}$$

It is also straightforward to show that $\mathbf{E}[r_j n_r(t)\,|\,\mathbf{s}_i] = 0$ for any t, $0 \leq t < T$. Thus, since r_j conditioned on $\mathbf{s}_i$ and $n_r(t)$ are Gaussian and uncorrelated, they are independent. Also, since the transmitted signal is independent of the noise, s_{ij} is independent of the process $n_r(t)$.

We now discuss the receiver design criterion and show that it is not affected by discarding $n_r(t)$. The goal of the receiver design is to minimize the probability of error in detecting the transmitted message m_i given received signal $r(t)$. In order to minimize $P_e = p(\hat{m} \neq m_i \mid r(t)) = 1 - p(\hat{m} = m_i \mid r(t))$, we maximize $p(\hat{m} = m_i \mid r(t))$. Therefore, the receiver output $\hat{m}$ given received signal $r(t)$ should correspond to the message m_i that maximizes $p(m_i \text{ sent} \mid r(t))$. Since there is a one-to-one mapping between messages and signal constellation points, this is equivalent to maximizing $p(\mathbf{s}_i \text{ sent} \mid r(t))$. Recalling that $r(t)$ is completely described by $\mathbf{r} = (r_1, \ldots, r_N)$ and $n_r(t)$, we have

$$
\begin{aligned}
p(\mathbf{s}_i \text{ sent} \mid r(t)) &= p((s_{i1}, \ldots, s_{iN}) \text{ sent} \mid (r_1, \ldots, r_N),\, n_r(t)) \\
&= \frac{p((s_{i1}, \ldots, s_{iN}) \text{ sent},\ (r_1, \ldots, r_N),\, n_r(t))}{p((r_1, \ldots, r_N),\, n_r(t))} \\
&= \frac{p((s_{i1}, \ldots, s_{iN}) \text{ sent},\ (r_1, \ldots, r_N))p(n_r(t))}{p(r_1, \ldots, r_N)p(n_r(t))} \\
&= p((s_{i1}, \ldots, s_{iN}) \text{ sent} \mid (r_1, \ldots, r_N)),
\end{aligned} \tag{5.26}
$$

where the third equality follows because $n_r(t)$ is independent of both $(r_1, \ldots, r_N)$ and $(s_{i1}, \ldots, s_{iN})$. This analysis shows that $\mathbf{r} = (r_1, \ldots, r_N)$ is a sufficient statistic for $r(t)$ in detecting m_i – in the sense that the probability of error is minimized by using only this sufficient statistic to estimate the transmitted signal and discarding the remainder noise. Since $\mathbf{r}$ is a sufficient statistic for the received signal $r(t)$, we call $\mathbf{r}$ the *received vector* associated with $r(t)$.

5.1.4 Decision Regions and the Maximum Likelihood Decision Criterion

We saw in the previous section that the optimal receiver minimizes error probability by selecting the detector output $\hat{m}$ that maximizes $1 - P_e = p(\hat{m} \text{ sent} \mid \mathbf{r})$. In other words, given a received vector $\mathbf{r}$, the optimal receiver selects $\hat{m} = m_i$ corresponding to the constellation $\mathbf{s}_i$ that satisfies $p(\mathbf{s}_i \text{ sent} \mid \mathbf{r}) \geq p(\mathbf{s}_j \text{ sent} \mid \mathbf{r})$ for all $j \neq i$. Let us define a set of *decision regions* $\{Z_1, \ldots, Z_M\}$ that are subsets of the signal space $\mathbb{R}^N$ by

$$
Z_i = \{\mathbf{r} \colon p(\mathbf{s}_i \text{ sent} \mid \mathbf{r}) > p(\mathbf{s}_j \text{ sent} \mid \mathbf{r})\ \forall j \neq i\}. \tag{5.27}
$$

Clearly these regions do not overlap. Moreover, they partition the signal space if there is no $\mathbf{r} \in \mathbb{R}^N$ for which $p(\mathbf{s}_i \text{ sent} \mid \mathbf{r}) = p(\mathbf{s}_j \text{ sent} \mid \mathbf{r})$ for some $j \neq i$. If such points exist then the signal space is partitioned into decision regions by arbitrarily assigning such points to decision region Z_i or Z_j. Once the signal space has been partitioned by decision regions, then for a received vector $\mathbf{r} \in Z_i$ the optimal receiver outputs the message estimate $\hat{m} = m_i$. Thus, the receiver processing consists of computing the received vector $\mathbf{r}$ from $r(t)$, finding which decision region Z_i contains $\mathbf{r}$, and outputting the corresponding message m_i. This process is illustrated in Figure 5.6, where we show a two-dimensional signal space with four decision regions $Z_1, \ldots, Z_4$ corresponding to four constellations $\mathbf{s}_1, \ldots, \mathbf{s}_4$. The received vector $\mathbf{r}$ lies in region Z_1, so the receiver will output the message m_1 as the best message estimate given received vector $\mathbf{r}$.

We now examine the decision regions in more detail. We will abbreviate $p(\mathbf{s}_i \text{ sent} \mid \mathbf{r} \text{ received})$ as $p(\mathbf{s}_i \mid \mathbf{r})$ and $p(\mathbf{s}_i \text{ sent})$ as $p(\mathbf{s}_i)$. By Bayes' rule,

$$
p(\mathbf{s}_i \mid \mathbf{r}) = \frac{p(\mathbf{r} \mid \mathbf{s}_i)p(\mathbf{s}_i)}{p(\mathbf{r})}. \tag{5.28}
$$

To minimize error probability, the receiver output $\hat{m} = m_i$ corresponds to the constellation $\mathbf{s}_i$ that maximizes $p(\mathbf{s}_i \mid \mathbf{r})$; that is, $\mathbf{s}_i$ must satisfy

$$
\arg\max_{\mathbf{s}_i} \frac{p(\mathbf{r} \mid \mathbf{s}_i)p(\mathbf{s}_i)}{p(\mathbf{r})} = \arg\max_{\mathbf{s}_i} p(\mathbf{r} \mid \mathbf{s}_i)p(\mathbf{s}_i), \quad i = 1, \ldots, M, \tag{5.29}
$$

where the second equality follows from the fact that $p(\mathbf{r})$ is not a function of $\mathbf{s}_i$. Assuming equally likely messages $(p(\mathbf{s}_i) = 1/M)$, the receiver output $\hat{m} = m_i$ corresponds to the constellation $\mathbf{s}_i$ that satisfies

$$
\arg\max_{\mathbf{s}_i} p(\mathbf{r} \mid \mathbf{s}_i), \quad i = 1, \ldots, M. \tag{5.30}
$$

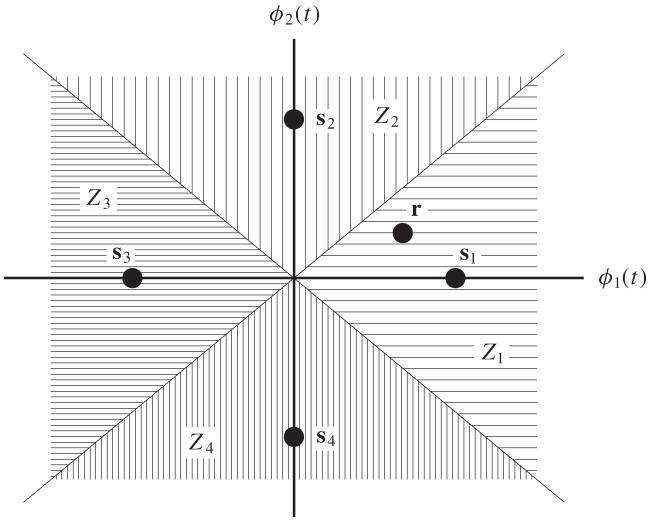


Figure 5.6: Decision regions.

Let us define the likelihood function associated with our receiver as

$$L(\mathbf{s}_i) = p(\mathbf{r} \mid \mathbf{s}_i). \tag{5.31}$$

Given a received vector $\mathbf{r}$, a *maximum likelihood receiver* outputs $\hat{m} = m_i$ corresponding to the constellation $\mathbf{s}_i$ that maximizes $L(\mathbf{s}_i)$. Since the log function is increasing in its argument, maximizing $L(\mathbf{s}_i)$ is equivalent to maximizing its log. Moreover, the constant factor $(\pi N_0)^{-N/2}$ in (5.25) does not affect the maximization of $L(\mathbf{s}_i)$ relative to $\mathbf{s}_i$. Thus, maximizing $L(\mathbf{s}_i)$ is equivalent to maximizing the *log likelihood function*, defined as $l(\mathbf{s}_i) = \ln[(\pi N_0)^{N/2} L(\mathbf{s}_i)]$. Using (5.25) for $L(\mathbf{s}_i) = p(\mathbf{r} \mid \mathbf{s}_i)$ then yields

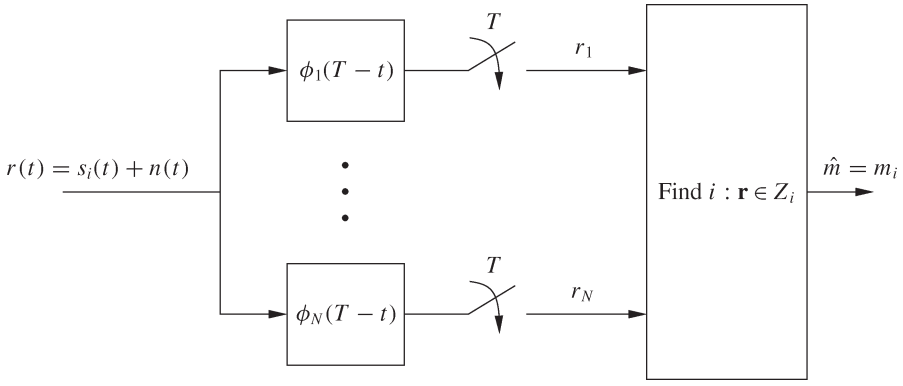


Figure 5.7: Matched filter receiver structure.

$$l(\mathbf{s}_i) = -\frac{1}{N_0} \sum_{j=1}^{N} (r_j - s_{ij})^2 = -\frac{1}{N_0} \|\mathbf{r} - \mathbf{s}_i\|^2. \tag{5.32}$$

Thus, the log likelihood function $l(\mathbf{s}_i)$ depends only on the distance between the received vector $\mathbf{r}$ and the constellation point $\mathbf{s}_i$. Moreover, from (5.32), $l(\mathbf{s}_i)$ is maximized by the constellation point $\mathbf{s}_i$ that is closest to the received vector $\mathbf{r}$.

The maximum likelihood receiver is implemented using the structure shown in Figure 5.4. First $\mathbf{r}$ is computed from $r(t)$, and then the signal constellation closest to $\mathbf{r}$ is determined as the constellation point $\mathbf{s}_i$ satisfying

$$\arg\min_{\mathbf{s}_i} \sum_{j=1}^{N} (r_j - s_{ij})^2 = \arg\min_{\mathbf{s}_i} \|\mathbf{r} - \mathbf{s}_i\|^2. \tag{5.33}$$

This $\mathbf{s}_i$ is determined from the decision region Z_i that contains $\mathbf{r}$, where Z_i is defined by

$$Z_i = \{\mathbf{r}\colon \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_j\|,\ \forall j = 1, \ldots, M,\ j \neq i\}, \quad i = 1, \ldots, M. \tag{5.34}$$

Finally, the estimated constellation $\mathbf{s}_i$ is mapped to the estimated message $\hat{m}$, which is output from the receiver. This result is intuitively satisfying, since the receiver decides that the transmitted constellation point is the one closest to the received vector. This maximum likelihood receiver structure is simple to implement because the decision criterion depends only on vector distances. This structure also minimizes the probability of message error at the receiver output when the transmitted messages are equally likely. However, if the messages and corresponding signal constellations are not equally likely then the maximum likelihood receiver does not minimize error probability; in order to minimize error probability, the decision regions Z_i must be modified to take into account the message probabilities, as indicated in (5.28).

An alternate receiver structure is shown in Figure 5.7. This structure makes use of a bank of filters matched to each of the different basis functions. We call a filter with impulse response $\psi(t) = \phi(T - t), 0 \leq t \leq T$, the *matched filter* to the signal $\phi(t)$, so Figure 5.7 is also called a *matched filter receiver*. It can be shown that if a given input signal is passed through a filter matched to that signal then the output SNR is maximized. One can also show that the sampled matched filter outputs $(r_1, \ldots, r_n)$ in Figure 5.7 are the same as the $(r_1, \ldots, r_n)$ in Figure 5.4, so the receivers depicted in these two figures are equivalent.

Example 5.2: For BPSK modulation, find decision regions Z_1 and Z_2 corresponding to constellations $s_1 = A$ and $s_2 = -A$ for $A > 0$.

Solution: The signal space is one-dimensional, so $\mathbf{r} = r \in \mathbb{R}$. By (5.34) the decision region $Z_1 \subset \mathbb{R}$ is defined by

$$Z_1 = \{r\colon \|r - A\| < \|r - (-A)\|\} = \{r\colon r > 0\}. \tag{5.35}$$

Thus, Z_1 contains all positive numbers on the real line. Similarly

$$Z_2 = \{r\colon \|r - (-A)\| < \|r - A\|\} = \{r\colon r < 0\}. \tag{5.36}$$

So Z_2 contains all negative numbers on the real line. For $r = 0$ the distance is the same to $s_1 = A$ and $s_2 = -A$, so we arbitrarily assign $r = 0$ to Z_2.

5.1.5 Error Probability and the Union Bound

We now analyze the error probability associated with the maximum likelihood receiver structure. For equally likely messages $p(m_i \text{ sent}) = 1/M$, we have

$$
\begin{aligned}
P_e &= \sum_{i=1}^{M} p(\mathbf{r} \notin Z_i \mid m_i \text{ sent}) p(m_i \text{ sent}) \\
&= \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r} \notin Z_i \mid m_i \text{ sent}) \\
&= 1 - \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r} \in Z_i \mid m_i \text{ sent}) \\
&= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} p(\mathbf{r} \mid m_i)\, d\mathbf{r} \\
&= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} p(\mathbf{r} = \mathbf{s}_i + \mathbf{n} \mid \mathbf{s}_i)\, d\mathbf{n} \\
&= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i - \mathbf{s}_i} p(\mathbf{n})\, d\mathbf{n}. \qquad (5.37)
\end{aligned}
$$

The integrals in (5.37) are over the N-dimensional subset $Z_i \subset \mathbb{R}^N$. We illustrate this error probability calculation in Figure 5.8, where the constellation points $\mathbf{s}_1, \ldots, \mathbf{s}_8$ are equally spaced around a circle with minimum separation $d_{\min}$. The probability of correct reception assuming the first symbol is sent, $p(\mathbf{r} \in Z_1 \mid m_1 \text{ sent})$, corresponds to the probability $p(\mathbf{r} = \mathbf{s}_1 + \mathbf{n} \in Z_1 \mid \mathbf{s}_1)$ that, when noise is added to the transmitted constellation $\mathbf{s}_1$, the resulting vector $\mathbf{r} = \mathbf{s}_1 + \mathbf{n}$ remains in the Z_1 region shown by the shaded area.

Figure 5.8 also indicates that the error probability is invariant to an angle rotation or axis shift of the signal constellation. The right side of the figure indicates a phase rotation of $\boldsymbol{\theta}$ and axis shift of $\mathbf{P}$ relative to the constellation on the left side. Thus, $\mathbf{s}_i' = \mathbf{s}_i e^{j\boldsymbol{\theta}} + \mathbf{P}$. The rotational invariance follows because the noise vector $\mathbf{n} = (n_1, \ldots, n_N)$ has components that are i.i.d. Gaussian random variables with zero mean; hence the polar representation $\mathbf{n} = |\mathbf{n}| e^{j\boldsymbol{\theta}}$ has $\boldsymbol{\theta}$ uniformly distributed, so the noise statistics are invariant to a phase rotation. The shift invariance follows from the fact that, if the constellation is shifted by some value $\mathbf{P} \in \mathbb{R}^N$, then the decision regions defined by (5.34) are also shifted by $\mathbf{P}$. Let $(\mathbf{s}_i, Z_i)$ denote a constellation point and corresponding decision region before the shift and $(\mathbf{s}_i', Z_i')$ the corresponding constellation point and decision region after the shift. It is then straightforward to show that $p(\mathbf{r} = \mathbf{s}_i + \mathbf{n} \in Z_i \mid \mathbf{s}_i) = p(\mathbf{r}' = \mathbf{s}_i' + \mathbf{n} \in Z_i' \mid \mathbf{s}_i')$. Thus, the error probability after an axis shift of the constellation points will remain unchanged.

Although (5.37) gives an exact solution to the probability of error, we cannot solve for this error probability in closed form. Therefore, we now investigate the union bound on error probability, which yields a closed-form expression that is a function of the distance between signal constellation points. Let A_{ik} denote the event that $\|\mathbf{r} - \mathbf{s}_k\| < \|\mathbf{r} - \mathbf{s}_i\|$ given that the constellation point $\mathbf{s}_i$ was sent. If the event A_{ik} occurs, then the constellation will be decoded in error because the transmitted constellation $\mathbf{s}_i$ is not the closest constellation point to the received vector $\mathbf{r}$. However, event A_{ik} does not necessarily imply that $\mathbf{s}_k$ will be decoded instead of $\mathbf{s}_i$, since there may be another constellation point $\mathbf{s}_l$ with $\|\mathbf{r} - \mathbf{s}_l\| < \|\mathbf{r} - \mathbf{s}_k\| < \|\mathbf{r} - \mathbf{s}_i\|$. The constellation is decoded correctly if $\|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_k\|$ for all $k \neq i$. Thus

$$
P_e(m_i \text{ sent}) = p\left(\bigcup_{\substack{k=1 \\ k \neq i}}^{M} A_{ik} \right) \leq \sum_{\substack{k=1 \\ k \neq i}}^{M} p(A_{ik}), \qquad (5.38)
$$

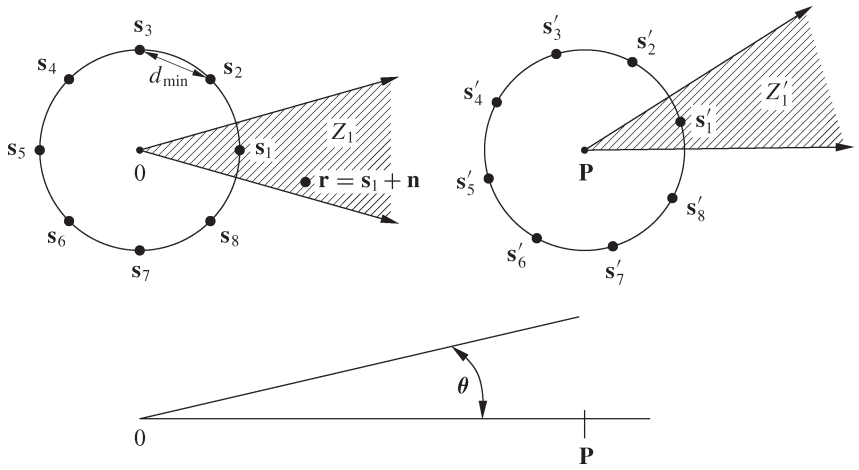


Figure 5.8: Error probability integral and its rotational/shift invariance.

where the inequality follows from the union bound on probability, defined below.

Let us now consider $p(A_{ik})$ more closely. We have

$$\begin{aligned} p(A_{ik}) &= p(\|\mathbf{s}_k - \mathbf{r}\| < \|\mathbf{s}_i - \mathbf{r}\| \mid \mathbf{s}_i \text{ sent}) \\ &= p(\|\mathbf{s}_k - (\mathbf{s}_i + \mathbf{n})\| < \|\mathbf{s}_i - (\mathbf{s}_i + \mathbf{n})\|) \\ &= p(\|\mathbf{n} + \mathbf{s}_i - \mathbf{s}_k\| < \|\mathbf{n}\|); \end{aligned} \tag{5.39}$$

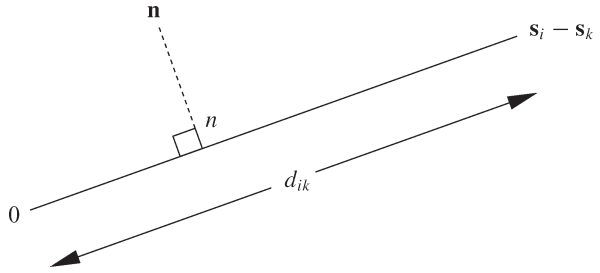


Figure 5.9: Noise projection.

that is, the probability of error equals the probability that the noise $\mathbf{n}$ is closer to the vector $\mathbf{s}_i - \mathbf{s}_k$ than to the origin. Recall that the noise has a mean of zero, so it is generally close to the origin. This probability does not depend on the entire noise component $\mathbf{n}$: it only depends on the projection of $\mathbf{n}$ onto the line connecting the origin and the point $\mathbf{s}_i - \mathbf{s}_k$, as shown in Figure 5.9. Given the properties of $\mathbf{n}$, the projection of $\mathbf{n}$ onto this one-dimensional line is a one-dimensional Gaussian random variable n with mean zero and variance $N_0/2$. The event A_{ik} occurs if n is closer to $\mathbf{s}_i - \mathbf{s}_k$ than to zero-that is, if $n > d_{ik}/2$, where $d_{ik} = \|\mathbf{s}_i - \mathbf{s}_k\|$ equals the distance between constellation points $\mathbf{s}_i$ and $\mathbf{s}_k$. Thus,

$$p(A_{ik}) = p\left(n > \frac{d_{ik}}{2}\right) = \int_{d_{ik}/2}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left[\frac{-v^2}{N_0}\right] dv = Q\left(\frac{d_{ik}}{\sqrt{2N_0}}\right). \tag{5.40}$$

Substituting (5.40) into (5.38) yields

$$P_e(m_i \text{ sent}) \leq \sum_{\substack{k=1 \\ k \neq i}}^{M} Q\left(\frac{d_{ik}}{\sqrt{2N_0}}\right), \tag{5.41}$$

where the Q-function, $Q(z)$, is defined as the probability that a Gaussian random variable X with mean 0 and variance 1 is greater than z:

$$Q(z) = p(X > z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \tag{5.42}$$

Summing (5.41) over all possible messages yields the *union bound*

$$P_e = \sum_{i=1}^{M} p(m_i) P_e(m_i \text{ sent}) \leq \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \neq i}}^{M} Q\left(\frac{d_{ik}}{\sqrt{2N_0}}\right). \tag{5.43}$$

Note that the Q-function cannot be solved for in closed form. It is related to the complementary error function as

$$Q(z) = \frac{1}{2} \text{ erfc}\left(\frac{z}{\sqrt{2}}\right). \tag{5.44}$$

We can also place an upper bound on $Q(z)$ with the closed-form expression

$$Q(z) \leq \frac{1}{z\sqrt{2\pi}} e^{-z^2/2}, \tag{5.45}$$

and this bound is tight for $z \gg 0$. A simpler upper bound can be obtained from the Chernoff bound [4] as

$$Q(z) \leq e^{-z^2/2}. \tag{5.46}$$

Defining the *minimum distance* of the constellation as $d_{\min} = \min_{i,k} d_{ik}$, we can simplify (5.43) with the looser bound

$$P_e \leq (M-1) Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right). \tag{5.47}$$

Using (5.45) for the Q-function yields a closed-form bound

$$P_e \leq \frac{M-1}{d_{\min}\sqrt{\pi/N_0}} \exp\left[\frac{-d_{\min}^2}{4N_0}\right]. \tag{5.48}$$

Finally, P_e is sometimes approximated as the probability of error associated with constellations at the minimum distance $d_{\min}$ multiplied by the number $M_{d_{\min}}$ of neighbors at this distance:

$$P_e \approx M_{d_{\min}} Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right). \tag{5.49}$$

This approximation is called the *nearest neighbor approximation* to P_e. When different constellation points have a different number of nearest neighbors or different minimum distances, the bound can be averaged over the bound associated with each constellation point. Note that the nearest neighbor approximation will always be less than the loose bound (5.47) since $M \geq M_{d_{\min}}$. It will also be slightly less than the union bound (5.43), since the nearest neighbor approximation does not include the error associated with constellations farther apart than the minimum distance. However, the nearest neighbor approximation is quite close to the exact probability of symbol error at

high SNRs, since for x and y large with $x > y$, $Q(x) \ll Q(y)$ owing to the exponential falloff of the Gaussian distribution in (5.42). This indicates that the probability of mistaking a constellation point for another point that is not one of its nearest neighbors is negligible at high SNRs. A rigorous derivation for (5.49) is made in [5] and also referenced in [6]. Moreover, [5] indicates that (5.49) captures the performance degradation due to imperfect receiver conditions such as slow carrier drift with an appropriate adjustment of the constants. The appeal of the nearest neighbor approximation is that it depends only on the minimum distance in the signal constellation and the number of nearest neighbors for points in the constellation.

Example 5.3: Consider a signal constellation in $\mathbb{R}^2$ defined by $s_1 = (A, 0)$, $s_2 = (0, A)$, $s_3 = (-A, 0)$, and $s_4 = (0, -A)$. Assume $A/\sqrt{N_0} = 4$. Find the minimum distance and the union bound (5.43), looser bound (5.47), closed-form bound (5.48), and nearest neighbor approximation (5.49) on P_e for this constellation set.
Solution: The constellation is as depicted in Figure 5.3 with the radius of the circle equal to A. By symmetry, we need only consider the error probability associated with one of the constellation points, since it will be the same for the others. We focus on the error associated with transmitting constellation point $\mathbf{s}_1$. The minimum distance to this constellation point is easily computed as $d_{\min} = d_{12} = d_{23} = d_{34} = d_{14} = \sqrt{A^2 + A^2} = \sqrt{2A^2}$. The distance to the other constellation points are $d_{13} = d_{24} = 2A$. By symmetry, $P_e(m_i \text{ sent}) = P_e(m_j \text{ sent})$ for $j \neq i$, so the union bound simplifies to

$$\begin{aligned} P_e &\leq \sum_{j=2}^{4} Q\left(\frac{d_{1j}}{\sqrt{2N_0}}\right) \\ &= 2Q\left(\frac{A}{\sqrt{N_0}}\right) + Q\left(\frac{\sqrt{2}A}{\sqrt{N_0}}\right) = 2Q(4) + Q(\sqrt{32}) = 3.1679 \cdot 10^{-5}. \end{aligned} \tag{5.50}$$

The looser bound yields

$$P_e \leq 3Q(4) = 9.5014 \cdot 10^{-5}, \tag{5.51}$$

which is roughly a factor of 3 looser than the union bound. The closed-form bound yields

$$P_e \leq \frac{3}{\sqrt{2\pi A^2/N_0}} \exp\left[\frac{-.5A^2}{N_0}\right] = 1.004 \cdot 10^{-4}, \tag{5.52}$$

which differs from the union bound by about an order of magnitude. Finally, the nearest neighbor approximation yields

$$P_e \approx 2Q(4) = 3.1671 \cdot 10^{-5}, \tag{5.53}$$

which (as expected) is approximately equal to the union bound.

Note that, for binary modulation (where $M = 2$), there is only one way to make an error and $d_{\min}$ is the distance between the two signal constellation points, so the bound (5.47) is exact:

$$P_b = Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right). \tag{5.54}$$

The square of the minimum distance $d_{\min}$ in (5.48) and (5.54) is typically proportional to the SNR of the received signal, as discussed in Chapter 6. Thus, error probability is reduced by increasing the received signal power.

Recall that P_e is the probability of a symbol (message) error: $P_e = p(\hat{m} \neq m_i \mid m_i \text{ sent})$, where m_i corresponds to a message with $\log_2 M$ bits. However, system designers are typically more interested in the bit

error probability (also called the bit error rate, BER) than in the symbol error probability, because bit errors drive the performance of higher-layer networking protocols and end-to-end performance. Thus, we would like to design the mapping of the M possible bit sequences to messages m_i $(i = 1, \ldots, M)$ so that a decoding error associated with an adjacent decision region, which is the most likely way to make an error, corresponds to only one bit error. With such a mapping – and with the assumption that mistaking a signal constellation point for a point other than one of its nearest neighbors has a very low probability – we can make the approximation

$$P_b \approx \frac{P_e}{\log_2 M}. \tag{5.55}$$

The most common form of mapping in which mistaking a constellation point for one of its nearest neighbors results in a single bit error is called Gray coding. Mapping by Gray coding is discussed in more detail in Section 5.3. Signal space concepts are applicable to any modulation where bits are encoded as one of several possible analog signals, including the amplitude, phase, and frequency modulations discussed in what follows.

5.2 Passband Modulation Principles

The basic principle of passband digital modulation is to encode an information bit stream into a carrier signal, which is then transmitted over a communications channel. Demodulation is the process of extracting this information bit stream from the received signal. Corruption of the transmitted signal by the channel can lead to bit errors in the demodulation process. The goal of modulation is to send bits at a high data rate while minimizing the probability of data corruption.

In general, modulated carrier signals encode information in the amplitude $\alpha(t)$, frequency $f(t)$, or phase $\theta(t)$ of a carrier signal. Thus, the modulated signal can be represented as

$$s(t) = \alpha(t)\cos[2\pi(f_c + f(t))t + \theta(t) + \phi_0] = \alpha(t)\cos(2\pi f_c t + \phi(t) + \phi_0), \tag{5.56}$$

where $\phi(t) = 2\pi f(t)t + \theta(t)$ and ϕ_0 is the phase offset of the carrier. This representation combines frequency and phase modulation into angle modulation.

We can rewrite the right-hand side of (5.56) in terms of its in-phase and quadrature components as

$$\begin{aligned} s(t) &= \alpha(t)\cos(\phi(t) + \phi_0)\cos(2\pi f_c t) - \alpha(t)\sin(\phi(t) + \phi_0)\sin(2\pi f_c t) \\ &= s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t), \end{aligned} \tag{5.57}$$

where $s_I(t) = \alpha(t)(\cos\phi(t) + \phi_0)$ is the in-phase component of $s(t)$ and where $s_Q(t) = \alpha(t)(\sin\phi(t) + \phi_0)$ is its quadrature component. We can also write $s(t)$ in terms of its equivalent lowpass representation as

$$s(t) = \mathrm{Re}\{u(t)e^{j2\pi f_c t}\}, \tag{5.58}$$

where $u(t) = s_I(t) + js_Q(t)$. This representation, described in more detail in Appendix A, is useful because receivers typically process the in-phase and quadrature signal components separately.

5.3 Amplitude and Phase Modulation

In amplitude and phase modulation, the information bit stream is encoded in the amplitude and/or phase of the transmitted signal. Specifically: over a time interval of T_s, $K = \log_2 M$ bits are encoded into the amplitude and/or phase of the transmitted signal $s(t)$, $0 \le t < T_s$. The transmitted signal over this period, $s(t) = s_I(t)\cos(2\pi f_c t) -$

$s_Q(t)\sin(2\pi f_c t)$, can be written in terms of its signal space representation as $s(t) = s_{i1}\phi_1(t) + s_{i2}\phi_2(t)$ with basis functions $\phi_1(t) = g(t)\cos(2\pi f_c t + \phi_0)$ and $\phi_2(t) = -g(t)\sin(2\pi f_c t + \phi_0)$, where $g(t)$ is a shaping pulse. For $\phi_0 = 0$, to send the ith message over the time interval $[kT, (k+1)T)$ we set $s_I(t) = s_{i1}g(t)$ and $s_Q(t) = s_{i2}g(t)$. These in-phase and quadrature signal components are baseband signals with spectral characteristics determined by the pulse shape $g(t)$. In particular, their bandwidth B equals the bandwidth of $g(t)$, and the transmitted signal $s(t)$ is a passband signal with center frequency f_c and passband bandwidth $2B$. In practice we take $B = K_g/T_s$, where K_g depends on the pulse shape: for rectangular pulses $K_g = .5$ and for raised cosine pulses $.5 \leq K_g \leq 1$, as discussed in Section 5.5. Thus, for rectangular pulses the bandwidth of $g(t)$ is. $5/T_s$ and the bandwidth of $s(t)$ is $1/T_s$. The signal constellation for amplitude and phase modulation is defined based on the constellation points $\{(s_{i1}, s_{i2}) \in \mathbb{R}^2, i = 1, \ldots, M\}$. The equivalent lowpass representation of $s(t)$ is

$$s(t) = \mathrm{Re}\{x(t)e^{j\phi_0}e^{j2\pi f_c t}\}, \tag{5.59}$$

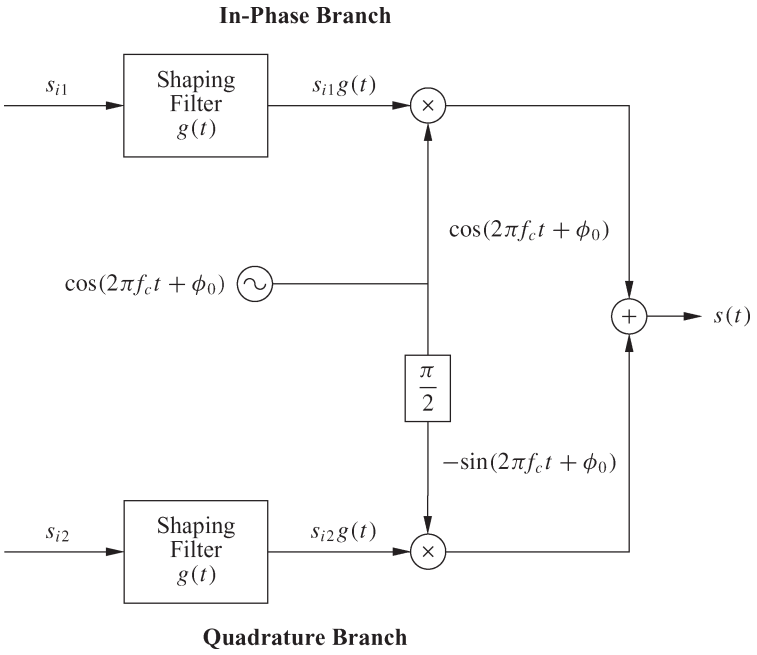


Figure 5.10: Amplitude/phase modulator.

where $x(t) = (s_{i1} + js_{i2})g(t)$. The constellation point $s_i = (s_{i1}, s_{i2})$ is called the *symbol* associated with the $\log_2 M$ bits, and T_s is called the *symbol time*. The bit rate for this modulation is K bits per symbol or $R = \log_2 M/T_s$ bits per second.

There are three main types of amplitude/phase modulation:

- pulse amplitude modulation (MPAM) – information encoded in amplitude only;
- phase-shift keying (MPSK) – information encoded in phase only;
- quadrature amplitude modulation (MQAM) – information encoded in both amplitude and phase.

The number of bits per symbol $K = \log_2 M$, the signal constellation $\{\mathbf{s}_i, i = 1, \ldots, M\}$, and the choice of pulse shape $g(t)$ determine the digital modulation design. The pulse shape $g(t)$ is chosen to improve spectral efficiency and combat ISI, as discussed in Section 5.5.

Amplitude and phase modulation over a given symbol period can be generated using the modulator structure shown in Figure 5.10. Note that the basis functions in this figure have an arbitrary phase ϕ_0 associated with the transmit oscillator. Demodulation over each symbol period is performed using the demodulation structure of Figure 5.11, which is equivalent to the structure of Figure 5.7 for $\phi_1(t) = g(t)\cos(2\pi f_c t + \phi)$ and $\phi_2(t) = -g(t)\sin(2\pi f_c t + \phi)$. Typically the receiver includes some additional circuitry for *carrier phase recovery* that matches the carrier phase ϕ at the receiver to the carrier phase ϕ_0 at the transmitter;[1] this is known as *coherent detection*. If $\phi - \phi_0 = \triangle\phi \neq 0$ then the in-phase branch will have an unwanted term associated with the quadrature branch and vice versa; that is, $r_1 = s_{i1}\cos(\triangle\phi) + s_{i2}\sin(\triangle\phi) + n_1$ and $r_2 = -s_{i1}\sin(\triangle\phi) + s_{i2}\cos(\triangle\phi) + n_2$, which can result in significant performance degradation. The receiver structure also assumes that the sampling function every T_s seconds is synchronized to the start of the symbol period, which is called *synchronization* or *timing recovery*. Receiver synchronization and carrier phase recovery are complex receiver operations that can be highly challenging in wireless environments. These operations are discussed in more detail in Section 5.6. We will assume perfect carrier recovery in our discussion of MPAM, MPSK, and MQAM and therefore set $\phi = \phi_0 = 0$ for their analysis.

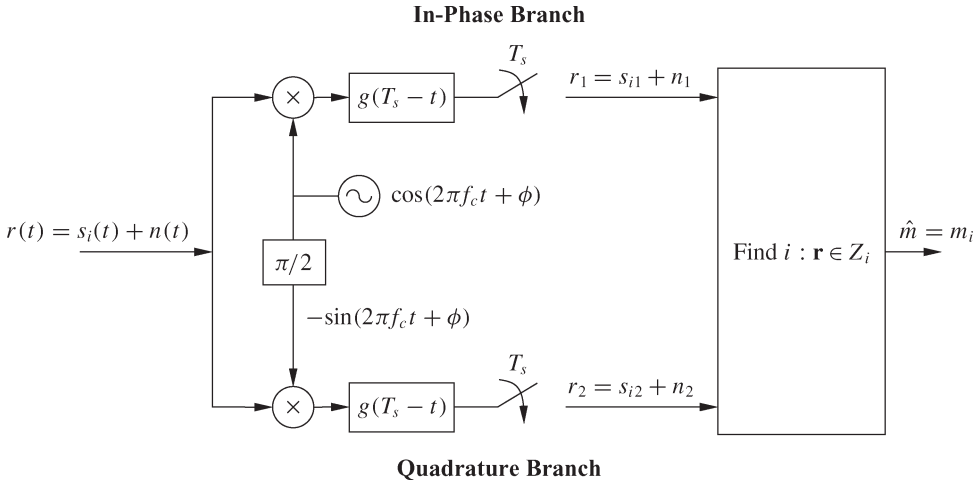


Figure 5.11: Amplitude/phase demodulator (coherent: $\phi = \phi_0$).

5.3.1 Pulse Amplitude Modulation (MPAM)

We will start by looking at the simplest form of linear modulation, one- dimensional MPAM, which has no quadrature component ($s_{i2} = 0$). For MPAM, all of the information is encoded into the signal amplitude A_i. The transmitted signal over one symbol time is given by

$$s_i(t) = \mathrm{Re}\{A_i g(t) e^{j2\pi f_c t}\} = A_i g(t)\cos(2\pi f_c t),\ 0 \leq t \leq T_s;\ T_s \gg 1/f_c, \tag{5.60}$$

where $A_i = (2i - 1 - M)d, i = 1, 2, \ldots, M$. The signal constellation is thus $\{A_i, i = 1, \ldots, M\}$, which is parameterized by the distance d. This distance is typically a function of the signal energy. The pulse shape $g(t)$ must satisfy (5.12) and (5.13). The minimum distance between constellation points is $d_{\min} = \min_{i,j}|A_i - A_j| = 2d$. The amplitude of the transmitted signal takes on M different values, which implies that each pulse conveys $\log_2 M = K$ bits per symbol time T_s.

[1] In fact, an additional phase term of $-2\pi f_c \tau$ will result from a propagation delay of τ in the channel. Thus, coherent detection requires the receiver phase $\phi = \phi_0 - 2\pi f_c \tau$, as discussed in more detail in Section 5.6.

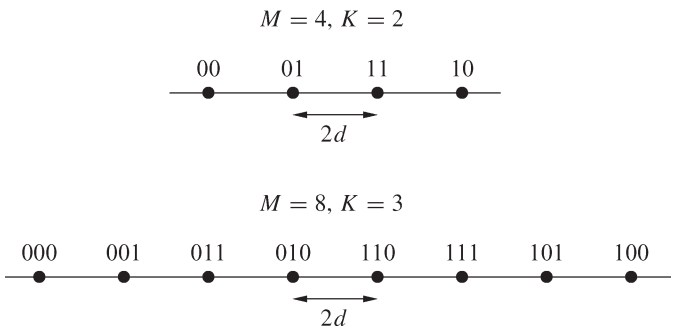


Figure 5.12: Gray encoding for MPAM.

Over each symbol period, the MPAM signal associated with the ith constellation has energy

$$E_{si} = \int_0^{T_s} s_i^2(t)\,dt = \int_0^{T_s} A_i^2 g^2(t)\cos^2(2\pi f_c t)\,dt = A_i^2, \tag{5.61}$$

since the pulse shape must satisfy (5.12).[2] Note that the energy is not the same for each signal $s_i(t), i = 1, \ldots, M$. Assuming equally likely symbols, the average energy is

$$\overline{E}_s = \frac{1}{M}\sum_{i=1}^{M} A_i^2. \tag{5.62}$$

The constellation mapping is usually done by Gray encoding, where the messages associated with signal amplitudes that are adjacent to each other differ by one bit value, as illustrated in Figure 5.12. With this encoding method, if noise causes the demodulation process to mistake one symbol for an adjacent one (the most likely type of error), the result is only a single bit error in the sequence of K bits. Gray codes can also be designed for MPSK and square MQAM constellations but not for rectangular MQAM.

Example 5.4: For $g(t) = \sqrt{2/T_s}(0 \le t < T_s)$ a rectangular pulse shape, find the average energy of 4-PAM modulation.
Solution: For 4-PAM the A_i values are $A_i = \{-3d, -d, d, 3d\}$, so the average energy is

$$\overline{E}_s = \frac{d^2}{4}(9 + 1 + 1 + 9) = 5d^2. \tag{5.63}$$

The decision regions $Z_i, i = 1, \ldots, M$, associated with pulse amplitude $A_i = (2i - 1 - M)d$ for $M = 4$ and $M = 8$ are shown in Figure 5.13. Mathematically, for any M these decision regions are defined by

$$Z_i = \begin{cases} (-\infty, A_i + d) & i = 1, \\ [A_i - d, A_i + d) & 2 \le i \le M - 1, \\ [A_i - d, \infty) & i = M. \end{cases} \tag{5.64}$$

From (5.60) we see that MPAM has only a single basis function $\phi_1(t) = g(t)\cos(2\pi f_c t)$. Thus, the coherent demodulator of Figure 5.11 for MPAM reduces to the demodulator shown in Figure 5.14, where the multithreshold device maps r to a decision region Z_i and outputs the corresponding bit sequence $\hat{m} = m_i = \{\mathrm{b}_1, \ldots, \mathrm{b}_K\}$.

[2] Recall from (5.8) that (5.12) and hence (5.61) are not exact equalities but rather very good approximations for $f_c T_s \gg 1$.

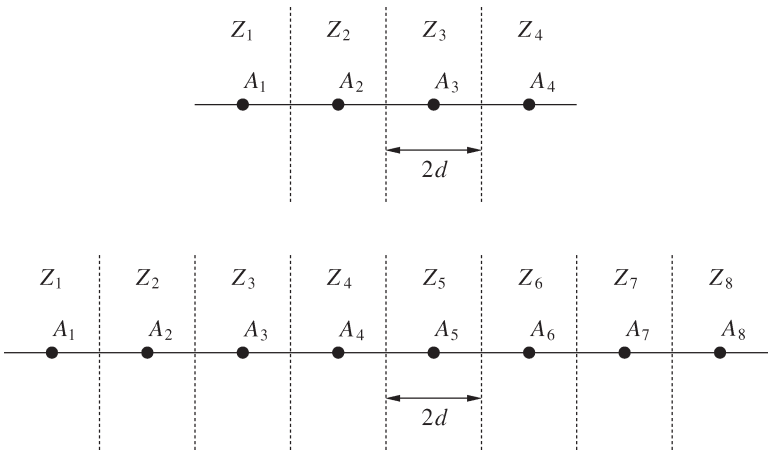


Figure 5.13: Decision regions for MPAM.

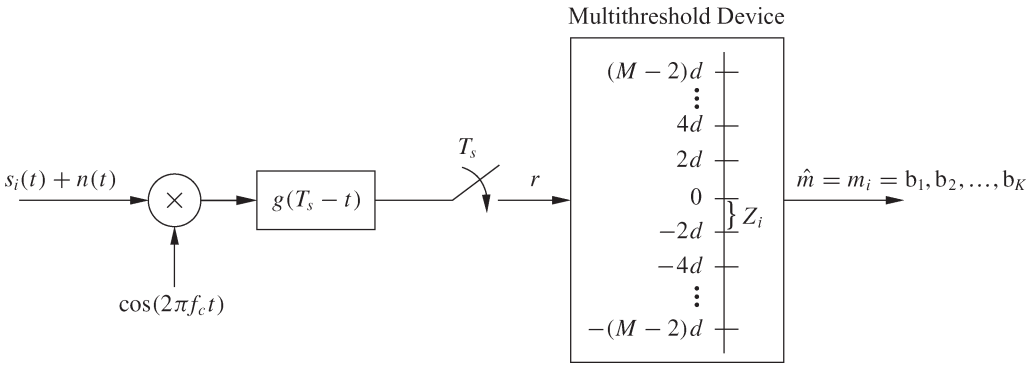


Figure 5.14: Coherent demodulator for MPAM.

5.3.2 Phase-Shift Keying (MPSK)

For MPSK, all of the information is encoded in the phase of the transmitted signal. Thus, the transmitted signal over one symbol time T_s is given by

$$\begin{aligned} s_i(t) &= \mathrm{Re}\{Ag(t)e^{j2\pi(i-1)/M}e^{j2\pi f_c t}\} \\ &= Ag(t)\cos\left[2\pi f_c t + \frac{2\pi(i-1)}{M}\right] \\ &= Ag(t)\cos\left[\frac{2\pi(i-1)}{M}\right]\cos 2\pi f_c t - Ag(t)\sin\left[\frac{2\pi(i-1)}{M}\right]\sin 2\pi f_c t \end{aligned} \tag{5.65}$$

for $0 \le t \le T_s$. Therefore, the constellation points or symbols (s_{i1}, s_{i2}) are given by $s_{i1} = A\cos[2\pi(i-1)/M]$ and $s_{i2} = A\sin[2\pi(i-1)/M]$ for $i = 1, \ldots, M$. The pulse shape $g(t)$ satisfies (5.12) and (5.13), and the $\theta_i = 2\pi(i-1)/M$ $(i = 1, 2, \ldots, M = 2^K)$ are the different phases in the signal constellation points that convey the information bits. The minimum distance between constellation points is $d_{\min} = 2A\sin(\pi/M)$, where A is typically a function of the signal energy. Note that 2-PSK is often referred to as binary PSK or BPSK, while 4-PSK is often called quadrature phase shift keying (QPSK) and is the same as MQAM with $M = 4$ (defined in Section 5.3.3).

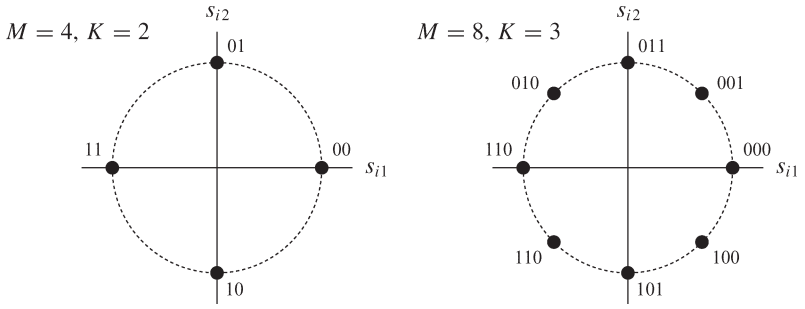


Figure 5.15: Gray encoding for MPSK.

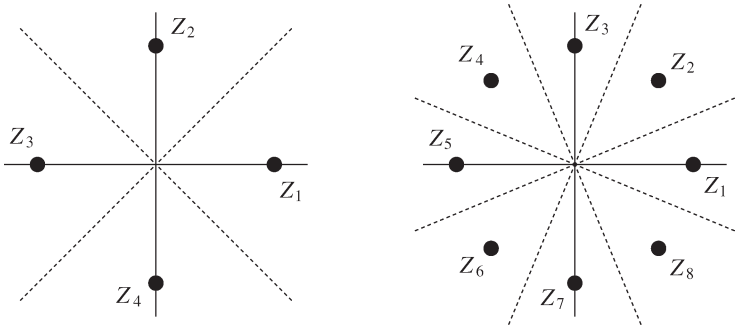


Figure 5.16: Decision regions for MPSK.

All possible transmitted signals $s_i(t)$ have equal energy:

$$E_{s_i} = \int_0^{T_s} s_i^2(t)dt = A^2. \tag{5.66}$$

Observe that for $g(t) = \sqrt{2/T_s}$ over a symbol time (i.e., a rectangular pulse) this signal has constant envelope, unlike the other amplitude modulation techniques MPAM and MQAM. However, rectangular pulses are spectrally inefficient, and more efficient pulse shapes give MPSK a nonconstant signal envelope. As for MPAM, constellation mapping is usually done by Gray encoding, where the messages associated with signal phases that are adjacent to each other differ by one bit value; see Figure 5.15. With this encoding method, mistaking a symbol for an adjacent one causes only a single bit error.

The decision regions $Z_i (i = 1, \ldots, M)$ associated with MPSK for $M = 4$ and $M = 8$ are shown in Figure 5.16. If we represent $\mathbf{r} = r_1 + jr_2 = re^{j\theta} \in \mathbb{R}^2$ in polar coordinates, then these decision regions for any M are defined by

$$Z_i = \{re^{j\theta} : 2\pi(i - 1.5)/M \leq \theta < 2\pi(i - .5)/M\}. \tag{5.67}$$

From (5.65) we see that MPSK has both in-phase and quadrature components, and thus the coherent demodulator is as shown in Figure 5.11. For the special case of BPSK, the decision regions as given in Example 5.2 simplify to $Z_1 = (r : r > 0)$ and $Z_2 = (r : r \leq 0)$. Moreover, BPSK has only a single basis function $\phi_1(t) = g(t)\cos(2\pi f_c t)$ and, since there is only a single bit transmitted per symbol time T_s, the bit time $T_b = T_s$. Thus, the coherent demodulator of Figure 5.11 for BPSK reduces to the demodulator shown in Figure 5.17, where the threshold device maps r to the positive or negative half of the real line and then outputs the corresponding bit value. We have assumed in this figure that the message corresponding to a bit value of $1, m_1 = 1$, is mapped to constellation

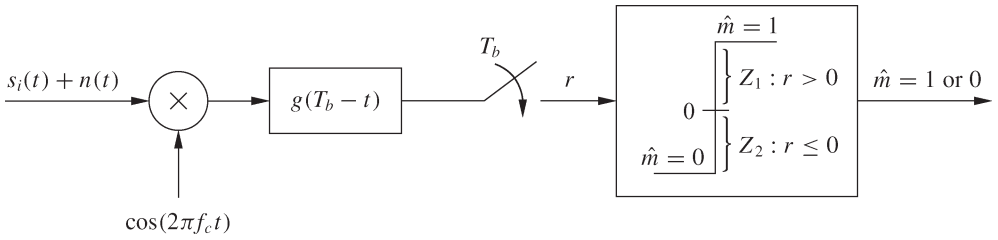


Figure 5.17: Coherent demodulator for BPSK.

point $s_1 = A$ and that the message corresponding to a bit value of $0, m_2 = 0$, is mapped to the constellation point $s_2 = -A$.

5.3.3 Quadrature Amplitude Modulation (MQAM)

For MQAM, the information bits are encoded in both the amplitude and phase of the transmitted signal. Thus, whereas both MPAM and MPSK have one degree of freedom in which to encode the information bits (amplitude or phase), MQAM has two degrees of freedom. As a result, MQAM is more spectrally efficient than MPAM and MPSK in that it can encode the most number of bits per symbol for a given average energy.

The transmitted signal is given by

$$\begin{aligned} s_i(t) &= \mathrm{Re}\{A_i e^{j\theta_i} g(t) e^{j2\pi f_c t}\} \\ &= A_i \cos(\theta_i) g(t) \cos(2\pi f_c t) - A_i \sin(\theta_i) g(t) \sin(2\pi f_c t),\ 0 \leq t \leq T_s, \end{aligned} \tag{5.68}$$

where the pulse shape $g(t)$ satisfies (5.12) and (5.13). The energy in $s_i(t)$ is

$$E_{s_i} = \int_0^{T_s} s_i^2(t) = A_i^2, \tag{5.69}$$

the same as for MPAM. The distance between any pair of symbols in the signal constellation is

$$d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\| = \sqrt{(s_{i1} - s_{j1})^2 + (s_{i2} - s_{j2})^2}. \tag{5.70}$$

For square signal constellations, where s_{i1} and s_{i2} take values on $(2i - 1 - L)d$ with $i = 1, 2, \ldots, L$, the minimum distance between signal points reduces to $d_{\min} = 2d$, the same as for MPAM. In fact, MQAM with square constellations of size L^2 is equivalent to MPAM modulation with constellations of size L on each of the in-phase and quadrature signal components. Common square constellations are 4-QAM and 16-QAM, which are shown in Figure 5.18. These square constellations have $M = L^2 = 2^{2l}$ constellation points, which are used to send $2l$ bits/symbol or l bits per dimension, where $l = .5 \log_2 M$. It can be shown that the average power of a square signal constellation with l bits per dimension, P_l, is proportional to $4^l/3$, and it follows that the average power for one more bit per dimension $P_{l+1} \approx 4P_l$. Thus, for square constellations it takes approximately 6 dB more power to send an additional 1 bit/dimension or 2 bits/symbol while maintaining the same minimum distance between constellation points.

Good constellation mappings can be hard to find for QAM signals, especially for irregular constellation shapes. In particular, it is hard to find a Gray code mapping where all adjacent symbols differ by a single bit. The decision regions $Z_i (i = 1, \ldots, M)$ associated with MQAM for $M = 16$ are shown in Figure 5.19. From

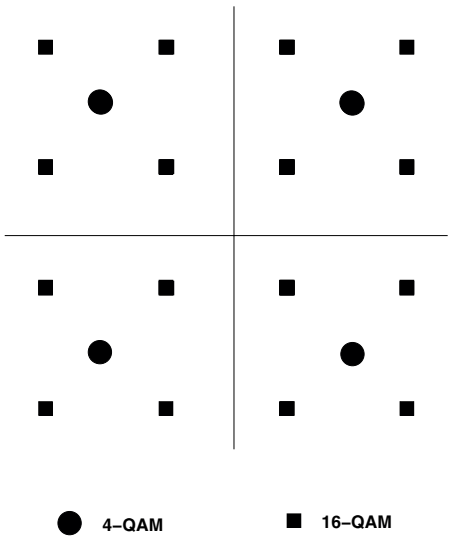


Figure 5.18: 4-QAM and 16-QAM constellations.

(5.68) we see that MQAM has both in-phase and quadrature components, and thus the coherent demodulator is as shown in Figure 5.11.

5.3.4 Differential Modulation

The information in MPSK and MQAM signals is carried in the signal phase. These modulation techniques therefore require coherent demodulation; that is, the phase of the transmitted signal carrier ϕ_0 must be matched to the phase of the receiver carrier ϕ. Techniques for phase recovery typically require more complexity and cost in the receiver, and they are also susceptible to phase drift of the carrier. Moreover, obtaining a coherent phase reference in a rapidly fading channel can be difficult. Issues associated with carrier phase recovery are discussed in more detail in Section 5.6. The difficulties as well as the cost and complexity associated with carrier phase recovery motivate the use of differential modulation techniques, which do not require a coherent phase reference at the receiver.

Differential modulation falls in the more general class of modulation with memory, where the symbol transmitted over time $[kT_s, (k+1)T_s)$ depends on the bits associated with the current message to be transmitted *and* on the bits transmitted over prior symbol times. The basic principle of differential modulation is to use the previous symbol as a phase reference for the current symbol, thus avoiding the need for a coherent phase reference at the receiver. Specifically, the information bits are encoded as the differential phase between the current symbol and the previous symbol. For example, in differential BPSK (referred to as DPSK), if the symbol over time $[(k-1)T_s, kT_s)$ has phase $\theta(k-1) = e^{j\theta_i}$ for $\theta_i = 0, \pi$, then to encode a 0-bit over $[kT_s, (k+1)T_s)$ the symbol would have phase $\theta(k) = e^{j\theta_i}$ and to encode a 1-bit the symbol would have phase $\theta(k) = e^{j(\theta_i+\pi)}$. In other words: a 0-bit is encoded by no change in phase, whereas a 1-bit is encoded as a phase change of π. Similarly, in 4-PSK modulation with differential encoding, the symbol phase over symbol interval $[kT_s, (k+1)T_s)$ depends on the current information bits over this time interval and the symbol phase over the previous symbol interval. The phase transitions for DQPSK modulation are summarized in Table 5.1.

Specifically, suppose the symbol over time $[(k-1)T_s, kT_s)$ has phase $\theta(k-1) = e^{j\theta_i}$. Then, over symbol time $[kT_s, (k+1)T_s)$, if the information bits are 00 then the corresponding symbol would have phase $\theta(k) = e^{j\theta_i}$; that is, to encode the bits 00, the symbol from symbol interval $[(k-1)T_s, kT_s)$ is repeated over the next interval $[kT_s, (k+1)T_s)$. If the two information bits to be sent at time interval $[kT_s, (k+1)T_s)$ are 01, then the corresponding symbol has phase $\theta(k) = e^{j(\theta_i+\pi/2)}$. For information bits 10 the symbol phase is $\theta(k) = e^{j(\theta_i-\pi/2)}$,

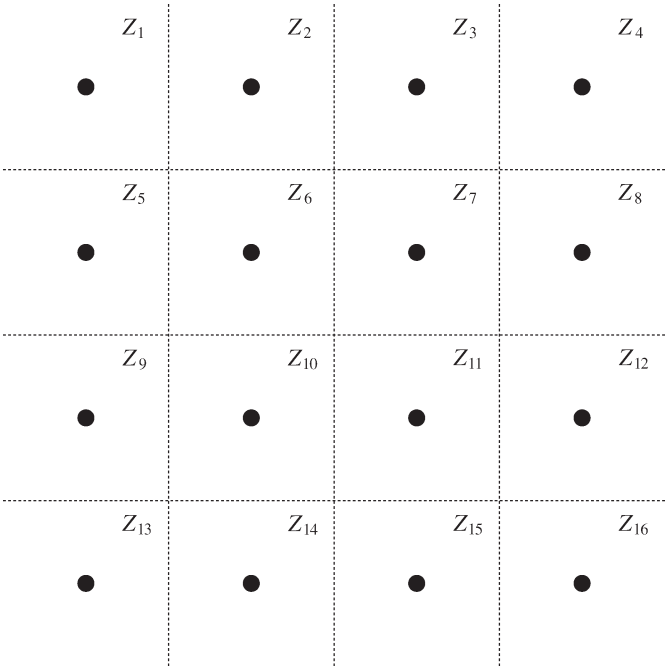


Figure 5.19: Decision regions for MQAM with $M = 16$.

Table 5.1: Mapping for DQPSK with Gray encoding

Bit sequence	**Phase transition**
00	0
01	$\pi/2$
10	$-\pi/2$
11	π

and for information bits 11 the symbol phase is $\theta(k) = e^{j(\theta_i+\pi)}$. We see that the symbol phase over symbol interval $[kT_s, (k+1)T_s)$ depends on the current information bits over this time interval and on the symbol phase θ_i over the previous symbol interval. Note that this mapping of bit sequences to phase transitions ensures that the most likely detection error – that of mistaking a received symbol for one of its nearest neighbors – results in a single bit error. For example, if the bit sequence 00 is encoded in the kth symbol then the kth symbol has the same phase as the $(k-1)$th symbol. Assume this phase is θ_i. The most likely detection error of the kth symbol is to decode it as one of its nearest neighbor symbols, which have phase $\theta_i \pm \pi/2$. But decoding the received symbol with phase $\theta_i \pm \pi/2$ would result in a decoded information sequence of either 01 or 10 – that is, it would differ by a single bit from the original sequence 00. More generally, we can use Gray encoding for the phase transitions in differential MPSK for any M, so that a message of all 0-bits results in no phase change, a message with a single 1-bit and the rest 0-bits results in the minimum phase change of $2\pi/M$, a message with two 1-bits and the rest 0-bits results in a phase change of $4\pi/M$, and so forth. Differential encoding is most common for MPSK signals, since the differential mapping is relatively simple. Differential encoding can also be done for MQAM with a more complex differential mapping. Differential encoding of MPSK is denoted by DMPSK, and for BPSK and QPSK by DPSK and DQPSK, respectively.

Example 5.5: Find the sequence of symbols transmitted using DPSK for the bit sequence 101110 starting at the kth symbol time, assuming the transmitted symbol at the $(k-1)$th symbol time was $\mathbf{s}(k-1) = Ae^{j\pi}$.

Solution: The first bit, a 1, results in a phase transition of π, so $\mathbf{s}(k) = A$. The next bit, a 0, results in no transition, so $\mathbf{s}(k+1) = A$. The next bit, a 1, results in another transition of π, so $\mathbf{s}(k+1) = Ae^{j\pi}$, and so on. The full symbol sequence corresponding to 101110 is $A, A, Ae^{j\pi}, A, Ae^{j\pi}, Ae^{j\pi}$.

The demodulator for differential modulation is shown in Figure 5.20. Assume the transmitted constellation at time k is $\mathbf{s}(k) = Ae^{j(\theta(k)+\phi_0)}$. Then the received vector associated with the sampler outputs is

$$\mathbf{r}(k) = r_1(k) + jr_2(k) = Ae^{j(\theta(k)+\phi_0)} + n(k), \tag{5.71}$$

where $n(k)$ is complex white Gaussian noise. The received vector at the previous time sample $k-1$ is thus

$$\mathbf{r}(k-1) = r_1(k-1) + jr_2(k-1) = Ae^{j(\theta(k-1)+\phi_0)} + n(k-1). \tag{5.72}$$

The phase difference between $\mathbf{r}(k)$ and $\mathbf{r}(k-1)$ determines which symbol was transmitted. Consider

$$\begin{aligned}\mathbf{r}(k)\mathbf{r}^*(k-1) =& A^2e^{j(\theta(k)-\theta(k-1))} + Ae^{j(\theta(k)+\phi_0)}n^*(k-1) \\ &+ Ae^{-j(\theta(k-1)+\phi_0)}n(k) + n(k)n^*(k-1).\end{aligned} \tag{5.73}$$

In the absence of noise $(n(k) = n(k-1) = 0)$ only the first term in (5.73) is nonzero, and this term yields the desired phase difference. The phase comparator in Figure 5.20 extracts this phase difference and outputs the corresponding symbol.

Differential modulation is less sensitive to a random drift in the carrier phase. However, if the channel has a nonzero Doppler frequency then the signal phase can decorrelate between symbol times, making the previous symbol a noisy phase reference. This decorrelation gives rise to an irreducible error floor for differential modulation over wireless channels with Doppler, as we shall discuss in Chapter 6.

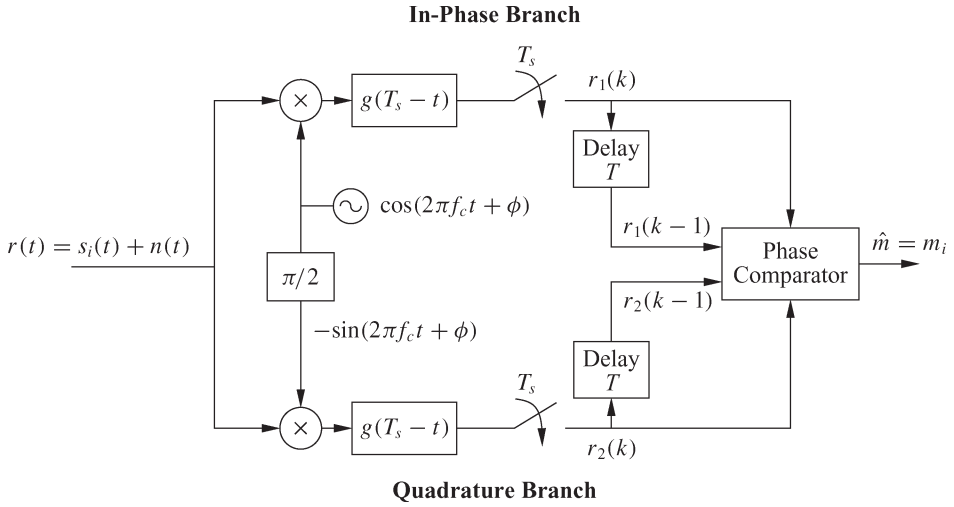


Figure 5.20: Differential PSK demodulator.

5.3.5 Constellation Shaping

Rectangular and hexagonal constellations have a better power efficiency than the square or circular constellations associated with MQAM and MPSK, respectively. These irregular constellations can save up to 1.3 dB of power at the expense of increased complexity in the constellation map [7]. The optimal constellation shape is a sphere in N-dimensional space, which must be mapped to a sequence of constellations in two-dimensional space in order to be generated by the modulator shown in Figure 5.10. The general conclusion in [7] is that, for uncoded modulation, the increased complexity of spherical constellations is not worth their energy gains, since coding can provide much better performance at less complexity cost. However, if a complex channel code is already being used and little further improvement can be obtained by a more complex code, constellation shaping may obtain around 1 dB of additional gain. An in-depth discussion of constellation shaping (and of constellations that allow a noninteger number of bits per symbol) can be found in [7].

5.3.6 Quadrature Offset

A linearly modulated signal with symbol $\mathbf{s}_i = (s_{i1}, s_{i2})$ will lie in one of the four quadrants of the signal space. At each symbol time kT_s the transition to a new symbol value in a different quadrant can cause a phase transition of up to 180°, which may cause the signal amplitude to transition through the zero point; these abrupt phase transitions and large amplitude variations can be distorted by nonlinear amplifiers and filters. The abrupt transitions are avoided by offsetting the quadrature branch pulse $g(t)$ by half a symbol period, as shown in Figure 5.21. This *quadrature offset* makes the signal less sensitive to distortion during symbol transitions.

Phase modulation with quadrature offset is usually abbreviated as OMPSK, where the O indicates the offset. For example, QPSK modulation with quadrature offset is referred to as OQPSK. Offset QPSK has the same spectral properties as QPSK for linear amplification, but it has higher spectral efficiency under nonlinear amplification because the maximum phase transition of the signal is 90°, corresponding to the maximum phase transition in either the in-phase or quadrature branch but not both simultaneously. Another technique to mitigate the amplitude fluctuations of a 180° phase shift used in the IS-136 standard for digital cellular is $\pi/4$-QPSK [8, 9]. This technique allows for a maximum phase transition of 135° degrees, versus 90° for offset QPSK and 180° for QPSK. Thus, $\pi/4$-QPSK has worse spectral properties than OQPSK under nonlinear amplification. However, $\pi/4$-QPSK can be differentially encoded to eliminate the need for a coherent phase reference, which is a significant advantage. Using differential encoding with $\pi/4$-QPSK is called $\pi/4$-DQPSK. The $\pi/4$-DQPSK modulation works as follows: the information bits are first differentially encoded as in DQPSK, which yields one of the four QPSK constellation points. Then, every other symbol transmission is shifted in phase by $\pi/4$. This periodic phase shift has a similar effect as the time offset in OQPSK: it reduces the amplitude fluctuations at symbol transitions, which makes the signal more robust against noise and fading.

5.4 Frequency Modulation

Frequency modulation encodes information bits into the frequency of the transmitted signal. Specifically: at each symbol time, $K = \log_2 M$ bits are encoded into the frequency of the transmitted signal $s(t)$, $0 \leq t < T_s$, resulting in a transmitted signal $s_i(t) = A\cos(2\pi f_i t + \phi_i)$, where i is the index of the ith message corresponding to the $\log_2 M$ bits and ϕ_i is the phase associated with the ith carrier. The signal space representation is $s_i(t) = \sum_j s_{ij}\phi_j(t)$, where $s_{ij} = A\delta(i-j)$ and $\phi_j(t) = \cos(2\pi f_j t + \phi_j)$, so the basis functions correspond to carriers at different frequencies and only one such basis function is transmitted in each symbol period. The basis functions are orthogonal for a minimum carrier frequency separation of $\triangle f = \min_{ij} |f_j - f_i| = .5/T_s$ for $\phi_i = \phi_j$ and of $\triangle f = 1/T_s$ for $\phi_i \neq \phi_j$.

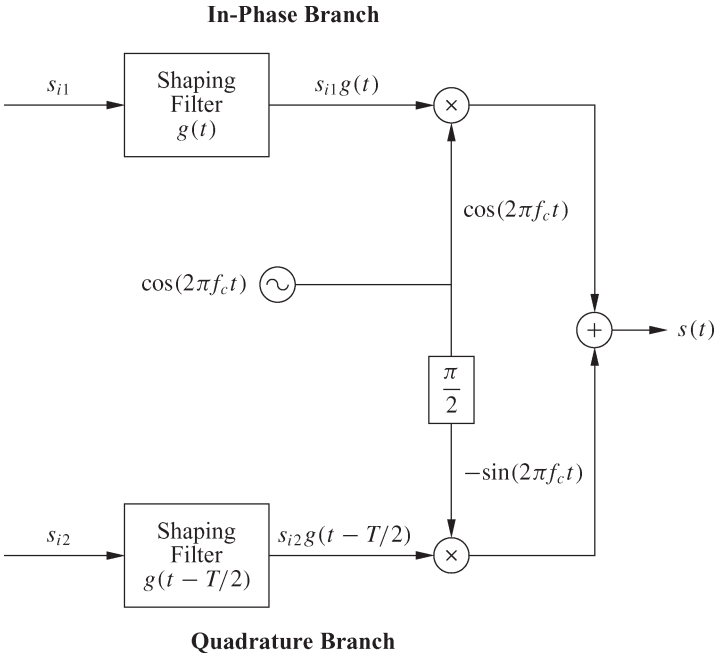


Figure 5.21: Modulator with quadrature offset.

Because frequency modulation encodes information in the signal frequency, the transmitted signal $s(t)$ has a constant envelope A. Since the signal is constant envelope, nonlinear amplifiers can be used with high power efficiency and hence the modulated signal is less sensitive to amplitude distortion introduced by the channel or the hardware. The price exacted for this robustness is a lower spectral efficiency: because the modulation technique is nonlinear, it tends to have a higher bandwidth occupancy than the amplitude and phase modulation techniques described in Section 5.3.

In its simplest form, frequency modulation over a given symbol period can be generated using the modulator structure shown in Figure 5.22. Demodulation over each symbol period is performed using the demodulation structure of Figure 5.23. Note that the demodulator of Figure 5.23 requires the jth carrier signal to be matched in phase to the jth carrier signal at the transmitter; this is similar to the coherent phase reference requirement in amplitude and phase modulation. An alternate receiver structure that does not require this coherent phase reference will be discussed in Section 5.4.3. Another issue in frequency modulation is that the different carriers shown in Figure 5.22 have different phases, $\phi_i \neq \phi_j$ for $i \neq j$, so at each symbol time T_s there will be a phase discontinuity in the transmitted signal. Such discontinuities can significantly increase signal bandwidth. Thus, in practice an alternate modulator is used that generates a frequency-modulated signal with continuous phase, as will be discussed in Section 5.4.2.

5.4.1 Frequency-Shift Keying (FSK) and Minimum-Shift Keying (MSK)

In MFSK the modulated signal is given by

$$s_i(t) = A\cos[2\pi f_c t + 2\pi\alpha_i \triangle f_c t + \phi_i],\ 0 \leq t < T_s, \tag{5.74}$$

where $\alpha_i = (2i - 1 - M)$ for $i = 1, 2, \ldots, M = 2^K$. The minimum frequency separation between FSK carriers is thus $2\triangle f_c$. MFSK consists of M basis functions $\phi_i(t) = \sqrt{2/T_s}\cos[2\pi f_c t + 2\pi\alpha_i \triangle f_c t + \phi_i],\ i = 1, \ldots, M$,

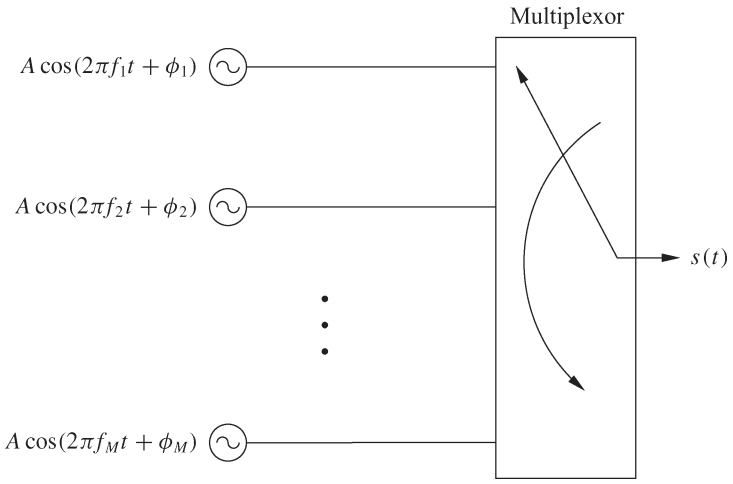


Figure 5.22: Frequency modulator.

where the $\sqrt{2/T_s}$ is a normalization factor to ensure that $\int_0^{T_s} \phi_i^2(t) = 1$. Over any given symbol time, only one basis function is transmitted through the channel.

A simple way to generate the MFSK signal is as shown in Figure 5.22, where M oscillators are operating at the different frequencies $f_i = f_c + \alpha_i \triangle f_c$ and the modulator switches between these different oscillators each symbol time T_s. However, this implementation entails a discontinuous phase transition at the switching times due to phase offsets between the oscillators, and this discontinuous phase leads to undesirable spectral broadening. (An FSK modulator that maintains continuous phase is discussed in the next section.) Coherent detection of MFSK uses the standard structure of Figure 5.23. For binary signaling the structure can be simplified to that shown in Figure 5.24, where the decision device outputs a 1-bit if its input is greater than zero and a 0-bit if its input is less than zero.

MSK is a special case of binary FSK where $\phi_1 = \phi_2$ and the frequency separation is $2\triangle f_c = .5/T_s$. Note that this is the minimum frequency separation that ensure $\langle s_i(t), s_j(t)\rangle = 0$ over a symbol time for $i \neq j$. Since signal orthogonality is required for demodulation, it follows that $2\triangle f_c = .5/T_s$ is the minimum possible frequency separation in FSK and so MSK is the minimum bandwidth FSK modulation.

5.4.2 Continuous-Phase FSK (CPFSK)

A better way to generate MFSK – one that eliminates the phase discontinuity – is to frequency modulate a single carrier with a modulating waveform, as in analog FM. In this case the modulated signal will be given by

$$s_i(t) = A\cos\left[2\pi f_c t + 2\pi\beta \int_{-\infty}^{t} u(\tau)\, d\tau\right] = A\cos[2\pi f_c t + \theta(t)], \tag{5.75}$$

where $u(t) = \sum_k a_k g(t - kT_s)$ is an MPAM signal modulated with the information bit stream, as described in Section 5.3.1. Clearly the phase $\theta(t)$ is continuous with this implementation. This form of MFSK is therefore called continuous-phase FSK, or CPFSK.

By Carson's rule [10], for β small the transmission bandwidth of $s(t)$ is approximately

$$B_s \approx 2M\triangle f_c + 2B_g, \tag{5.76}$$

where B_g is the bandwidth of the pulse shape $g(t)$ used in the MPAM modulating signal $u(t)$. By comparison, the bandwidth of a linearly modulated waveform with pulse shape $g(t)$ is roughly $B_s \approx 2B_g$. Thus, the spectral

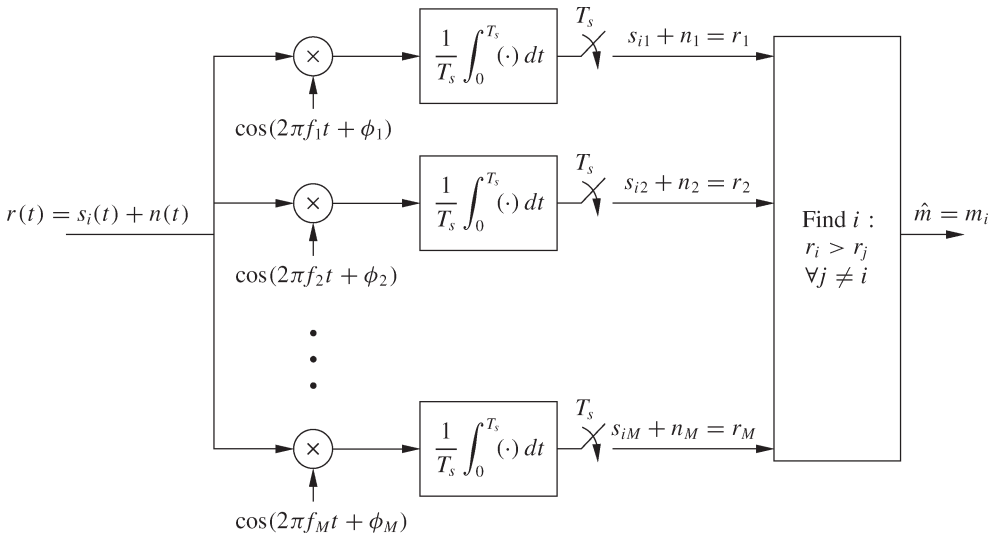


Figure 5.23: Frequency demodulator (coherent).

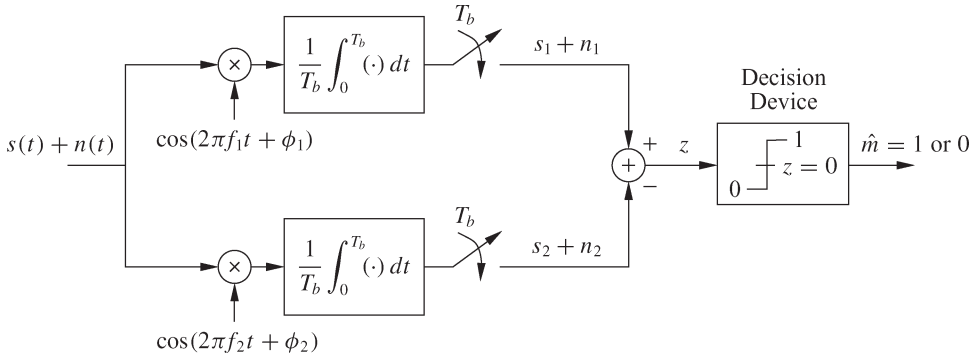


Figure 5.24: Demodulator for binary FSK.

occupancy of a CPFSK-modulated signal is larger than that of a linearly modulated signal by $M\triangle f_c \geq .5M/T_s$. The spectral efficiency penalty of CPFSK relative to linear modulation increases with data rate, in particular with the number of bits per symbol $K = \log_2 M$ and with the symbol rate $R_s = 1/T_s$.

Coherent detection of CPFSK can be done symbol-by-symbol or over a sequence of symbols. The sequence estimator is the optimal detector, since a given symbol depends on previously transmitted symbols and so it is optimal to detect (or estimate) all symbols simultaneously. However, sequence estimation can be impractical owing to the memory and computational requirements associated with making decisions based on sequences of symbols. Details on detectors for coherent demodulation of CPFSK can be found in [11, Chap. 4.9].

5.4.3 Noncoherent Detection of FSK

The receiver requirement for a coherent phase reference associated with each FSK carrier can be difficult and expensive to meet. The need for a coherent phase reference can be eliminated if the receiver first detects the energy of the signal at each frequency and, if the ith branch has the highest energy of all branches, then outputs message m_i. The modified receiver is shown in Figure 5.25.

Suppose the transmitted signal corresponds to frequency f_i:

$$
\begin{aligned}
s(t) &= A\cos(2\pi f_i t + \phi_i) \\
&= A\cos(\phi_i)\cos(2\pi f_i t) - A\sin(\phi_i)\sin(2\pi f_i t),\ 0 \le t < T_s.
\end{aligned} \tag{5.77}
$$

Let the phase ϕ_i represent the phase offset between the transmitter and receiver oscillators at frequency f_i. A coherent receiver with carrier signal $\cos(2\pi f_i t)$ detects only the first term $A\cos(\phi_i)\cos(2\pi f_i t)$ associated with the received signal, which can be close to zero for a phase offset $\phi_i \approx \pm\pi/2$. To get around this problem, in Figure 5.25 the receiver splits the received signal into M branches corresponding to each frequency $f_j,\ j = 1, \ldots, M$. For each such carrier frequency f_j, the received signal is multiplied by a noncoherent in-phase and quadrature carrier at that frequency, integrated over a symbol time, sampled, and then squared. For the jth branch the squarer output associated with the in-phase component is denoted as $A_{jI} + n_{jI}$ and the corresponding output associated with the quadrature component is denoted as $A_{jQ} + n_{jQ}$, where n_{jI} and n_{jQ} are due to the noise $n(t)$ at the receiver input. Then, if $i = j$, we have $A_{jI} = A^2\cos^2(\phi_i)$ and $A_{jQ} = A^2\sin^2(\phi_i)$; if $i \neq j$ then $A_{jI} = A_{jQ} = 0$. In the absence of noise, the input to the decision device of the ith branch will be $A^2\cos^2(\phi_i) + A^2\sin^2(\phi_i) = A^2$, independent of ϕ_i, and all other branches will have an input of zero. Thus, over each symbol period, the decision device outputs the bit sequence corresponding to frequency f_j if the jth branch has the largest input to the decision device. Note that the noncoherent receiver of Figure 5.25 still requires accurate synchronization for sampling. Synchronization issues are discussed in Section 5.6.

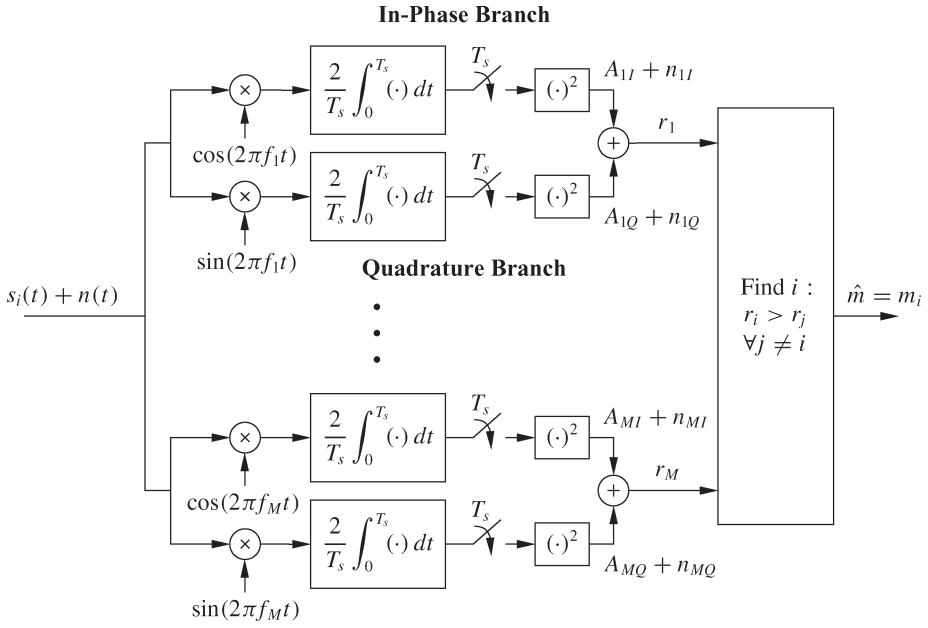


Figure 5.25: Noncoherent FSK demodulator.

5.5 Pulse Shaping

For amplitude and phase modulation, the bandwidth of the baseband and passband modulated signal is a function of the bandwidth of the pulse shape $g(t)$. If $g(t)$ is a rectangular pulse of width T_s, then the envelope of the signal is constant. However, a rectangular pulse has high spectral sidelobes, which can cause adjacent channel interference. Pulse shaping is a method for reducing sidelobe energy relative to a rectangular pulse; however, the shaping must

be done in such a way that intersymbol interference between pulses in the received signal is not introduced. Note that – prior to sampling the received signal – the transmitted pulse $g(t)$ is convolved with the channel impulse response $c(t)$ and the matched filter $g^*(-t)$; hence, in order to eliminate ISI prior to sampling, we must ensure that the effective received pulse $p(t) = g(t) * c(t) * g^*(-t)$ has no ISI. Since the channel model is AWGN, we assume $c(t) = \delta(t)$ so $p(t) = g(t) * g^*(-t)$ (in Chapter 11we will analyze ISI for more general channel impulse responses $c(t)$). To avoid ISI between samples of the received pulses, the effective pulse shape $p(t)$ must satisfy the *Nyquist criterion*, which requires the pulse to equal zero at the ideal sampling point associated with past or future symbols:

$$p(kT_s) = \begin{cases} p_0 = p(0) & k = 0, \\ 0 & k \neq 0. \end{cases} \tag{5.78}$$

In the frequency domain this translates to

$$\sum_{l=-\infty}^{\infty} P\left(f + \frac{l}{T_s}\right) = p_0 T_s. \tag{5.79}$$

The following pulse shapes all satisfy the Nyquist criterion.

1. *Rectangular pulses:* $g(t) = \sqrt{2/T_s}(0 \leq t \leq T_s)$, which yields the triangular effective pulse shape

$$p(t) = \begin{cases} 2 + 2t/T_s & -T_s \leq t < 0, \\ 2 - 2t/T_s & 0 \leq t < T_s, \\ 0 & \text{else.} \end{cases} \tag{5.80}$$

This pulse shape leads to constant envelope signals in MPSK but has poor spectral properties as a result of its high sidelobes.

2. *Cosine pulses:* $p(t) = \sin \pi t/T_s, 0 \leq t \leq T_s$. Cosine pulses are mostly used in OMPSK modulation, where the quadrature branch of the modulation is shifted in time by $T_s/2$. This leads to a constant amplitude modulation with sidelobe energy that is 10 dB lower than that of rectangular pulses.

3. *Raised cosine pulses:* These pulses are designed in the frequency domain according to the desired spectral properties. Thus, the pulse $p(t)$ is first specified relative to its Fourier transform:

$$P(f) = \begin{cases} T_s & 0 \leq |f| \leq \frac{1-\beta}{2T_s}, \\ \frac{T_s}{2}\left[1 - \sin\frac{\pi T_s}{\beta}\left(f - \frac{1}{2T_s}\right)\right] & \frac{1-\beta}{2T_s} \leq |f| \leq \frac{1+\beta}{2T_s}; \end{cases} \tag{5.81}$$

here β is defined as the rolloff factor, which determines the rate of spectral rolloff (see Figure 5.26). Setting $\beta = 0$ yields a rectangular pulse. The pulse $p(t)$ in the time domain corresponding to $P(f)$ is

$$p(t) = \frac{\sin(\pi t/T_s)}{\pi t/T_s} \frac{\cos(\beta \pi t/T_s)}{1 - 4\beta^2 t^2/T_s^2}. \tag{5.82}$$

The frequency- and time-domain properties of the raised cosine pulse are shown in Figures 5.26 and 5.27, respectively. The tails of this pulse in the time domain decay as $1/t^3$ (faster than for the previous pulse shapes), so a mistiming error in sampling leads to a series of inter-symbol interference components that converge. A variation of the raised cosine pulse is the root cosine pulse, derived by taking the square root of the frequency response for the raised cosine pulse. The root cosine pulse has better spectral properties than the raised cosine pulse but decays

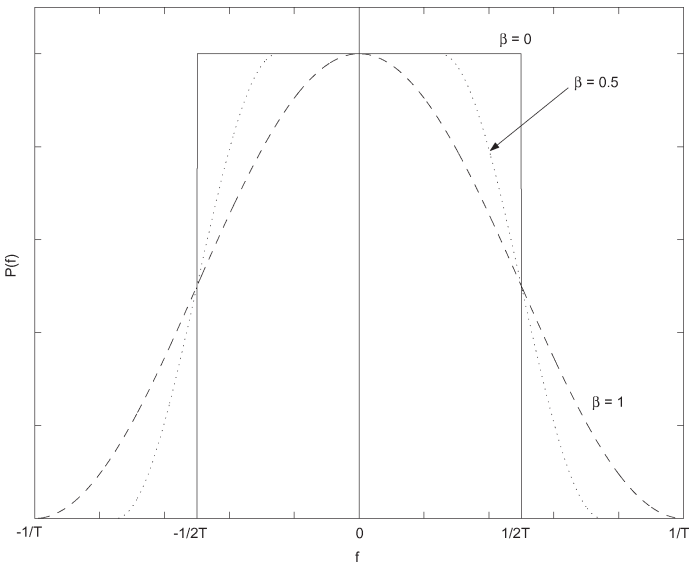


Figure 5.26: Frequency-domain (spectral) properties of the raised cosine pulse ($T = T_s$).

less rapidly in the time domain, which makes performance degradation due to synchronization errors more severe. Specifically, a mistiming error in sampling leads to a series of ISI components that may diverge.

Pulse shaping is also used with CPFSK to improve spectral efficiency, specifically in the MPAM signal that is frequency modulated to form the FSK signal. The most common pulse shape used in CPFSK is the Gaussian pulse shape, defined as

$$g(t) = \frac{\sqrt{\pi}}{\alpha} e^{-\pi^2 t^2/\alpha^2}, \tag{5.83}$$

where α is a parameter that dictates spectral efficiency. The spectrum of $g(t)$, which dictates the spectrum of the CPFSK signal, is given by

$$G(f) = e^{-\alpha^2 f^2}. \tag{5.84}$$

The parameter α is related to the 3-dB bandwidth of $g(t)$, B_g, by

$$\alpha = \frac{.5887}{B_g}. \tag{5.85}$$

Clearly, increasing α results in a higher spectral efficiency.

When the Gaussian pulse shape is applied to MSK modulation, it is abbreviated as GMSK. In general, GMSK signals have a high power efficiency (since they have a constant amplitude) and a high spectral efficiency (since the Gaussian pulse shape has good spectral properties for large α). For these reasons, GMSK is used in the GSM standard for digital cellular systems. Although this is a good choice for voice modulation, it is not the best choice for data. The Gaussian pulse shape does not satisfy the Nyquist criterion and so the pulse shape introduces ISI, which increases as α increases. Thus, improving spectral efficiency by increasing α leads to a higher ISI level, thereby creating an irreducible error floor from this self-interference. Since the required BER for voice is a relatively high $P_b \approx 10^{-3}$, the ISI can be fairly high and still maintain this target BER. In fact, it is generally used as a rule of thumb that $B_g T_s = .5$ is a tolerable amount of ISI for voice transmission with GMSK. However, a much lower BER is required for data, which will put more stringent constraints on the maximum α and corresponding

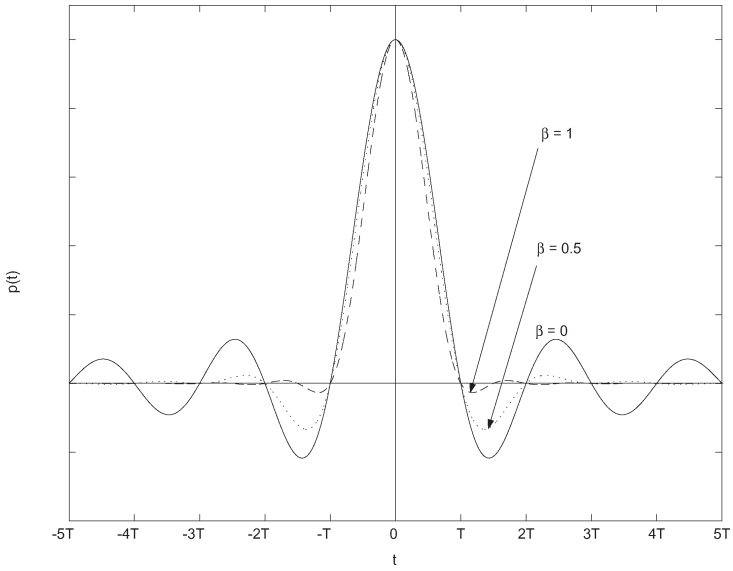


Figure 5.27: Time-domain properties of the raised cosine pulse ($T = T_s$).

minimum B_g, thereby decreasing the spectral efficiency of GMSK for data transmission. Techniques such as equalization can be used to mitigate the ISI in this case so that a tolerable BER is possible without significantly compromising spectral efficiency. However, it is more common to use linear modulation for spectrally efficient data transmission. Indeed, the data enhancements to GSM use linear modulation.

5.6 Symbol Synchronization and Carrier Phase Recovery

One of the most challenging tasks of a digital demodulator is to acquire accurate symbol timing and carrier phase information. Timing information, obtained via synchronization, is needed to delineate the received signal associated with a given symbol. In particular, timing information is used to drive the sampling devices associated with the demodulators for amplitude, phase, and frequency demodulation shown in Figures 5.11 and 5.23. Carrier phase information is needed in all coherent demodulators for both amplitude/phase and frequency modulation, as discussed in Sections 5.3 and 5.4.

This section gives a brief overview of standard techniques for synchronization and carrier phase recovery in AWGN channels. In this context the estimation of symbol timing and carrier phase falls under the broader category of signal parameter estimation in noise. Estimation theory provides the theoretical framework for studying this problem and for developing the maximum likelihood estimator of the carrier phase and symbol timing. However, most wireless channels suffer from time-varying multipath in addition to AWGN. Synchronization and carrier phase recovery is particularly challenging in such channels because multipath and time variations can make it extremely difficult to estimate signal parameters prior to demodulation. Moreover, there is little theory addressing good methods for estimation of carrier phase and symbol timing when these parameters are corrupted by time-varying multipath in addition to noise. In most performance analysis of wireless communication systems it is assumed that the receiver synchronizes to the multipath component with delay equal to the average delay spread;[3] then the channel is treated as AWGN for recovery of timing information and carrier phase. In practice, however, the receiver will synchronize to either the strongest multipath component or the first multipath component that exceeds a given power threshold. The other multipath components will then compromise the receiver's ability

[3]That is why delay spread is typically characterized by its rms value about its mean, as discussed in more detail in Chapter 2.

to acquire timing and carrier phase, especially in wideband systems like UWB. Multicarrier and spread-spectrum systems have additional considerations related to synchronization and carrier recovery, which will be discussed in Chapters 12 and 13.

The importance of synchronization and carrier phase estimation cannot be overstated: without them, wireless systems could not function. Moreover, as data rates increase and channels become more complex by adding additional degrees of freedom (e.g., multiple antennas), the tasks of receiver synchronization and phase recovery become even more complex and challenging. Techniques for synchronization and carrier recovery have been developed and analyzed extensively for many years, and they are continually evolving to meet the challenges associated with higher data rates, new system requirements, and more challenging channel characteristics. We give only a brief introduction to synchronization and carrier phase recovery techniques in this section. Comprehensive coverage of this topic and performance analysis of these techniques can be found in [12, 13]; more condensed treatments can be found in [11, Chap. 5] and [14].

5.6.1 Receiver Structure with Phase and Timing Recovery

The carrier phase and timing recovery circuitry for the amplitude and phase demodulator is shown in Figure 5.28. For BPSK only the in-phase branch of this demodulator is needed. For the coherent frequency demodulator of Figure 5.23, a carrier phase recovery circuit is needed for *each* of the distinct M carriers; the resulting circuit complexity motivates the need for noncoherent demodulators as described in Section 5.4.3. We see in Figure 5.28 that the carrier phase and timing recovery circuits operate directly on the received signal prior to demodulation.

Assuming an AWGN channel, the received signal $r(t)$ is a delayed version of the transmitted signal $s(t)$ plus AWGN $n(t)$: $r(t) = s(t-\tau) + n(t)$, where τ is the random propagation delay. Using the equivalent lowpass form we have $s(t) = \mathrm{Re}\{x(t)e^{j\phi_0}e^{j2\pi f_c t}\}$ and thus

$$r(t) = \mathrm{Re}\left\{x(t-\tau)e^{j\phi}e^{j2\pi f_c t}\right\} + n(t), \tag{5.86}$$

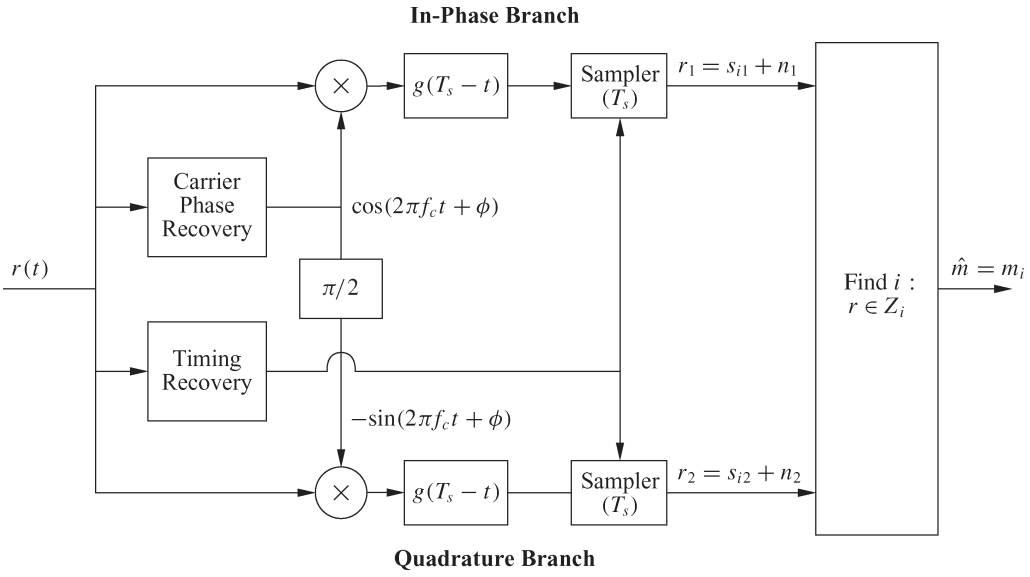


Figure 5.28: Receiver structure with carrier and timing recovery.

where $\phi = \phi_0 - 2\pi f_c \tau$ results from the transmit carrier phase and the propagation delay. Estimation of τ is needed for symbol timing, and estimation of ϕ is needed for carrier phase recovery. Let us express these two unknown

parameters as a vector $\psi = (\phi, \tau)$. Then we can express the received signal in terms of ψ as

$$r(t) = s(t; \psi) + n(t). \tag{5.87}$$

Parameter estimation must take place over some finite time interval $T_0 \geq T_s$. We call T_0 the *observation interval*. In practice, however, parameter estimation is initially done over this interval and is thereafter performed continually by updating the initial estimate using tracking loops. Our development here focuses on the initial parameter estimation over T_0. Discussion of parameter tracking can be found in [12, 13].

There are two common estimation methods for signal parameters in noise: the maximum likelihood (ML) criterion discussed in Section 5.1.4 in the context of receiver design, and the maximum a posteriori (MAP) criterion. The ML criterion chooses the estimate $\hat{\psi}$ that maximizes $p(r(t)|\psi)$ over the observation interval T_0, whereas the MAP criterion assumes some probability distribution $p(\psi)$ on ψ and then chooses the estimate $\hat{\psi}$ that maximizes

$$p(\psi|r(t)) = \frac{p(r(t)|\psi)p(\psi)}{p(r(t))} \tag{5.88}$$

over T_0. We assume that there is no prior knowledge of ψ, so that $p(\psi)$ becomes uniform and hence the MAP and ML criteria are equivalent.

To characterize the distribution $p(r(t)|\psi)$, $0 \leq t < T_0$, let us expand $r(t)$ over the observation interval along a set of orthonormal basis functions $\{\phi_k(t)\}$ as

$$r(t) = \sum_{k=1}^{K} r_k \phi_k(t), \quad 0 \leq t < T_0. \tag{5.89}$$

Because $n(t)$ is white with zero mean and power spectral density $N_0/2$, the distribution of the vector $\mathbf{r} = (r_1, \ldots, r_K)$ conditioned on the unknown parameter ψ is given by

$$p(\mathbf{r}|\psi) = \left(\frac{1}{\sqrt{\pi N_0}}\right)^K \exp\left[-\sum_{k=1}^{K} \frac{(r_k - s_k(\psi))^2}{N_0}\right], \tag{5.90}$$

where (by the basis expansion)

$$r_k = \int_{T_0} r(t)\phi_k(t)\, dt \tag{5.91}$$

and

$$s_k(\psi) = \int_{T_0} s(t; \psi)\phi_k(t)\, dt. \tag{5.92}$$

From these basis expansions we can show that

$$\sum_{k=1}^{K} [r_k - s_k(\psi)]^2 = \int_{T_0} [r(t) - s(t; \psi)]^2\, dt. \tag{5.93}$$

Using this in (5.90) yields that maximizing $p(\mathbf{r}|\psi)$ is equivalent to maximizing the *likelihood function*

$$\Lambda(\psi) = \exp\left[-\frac{1}{N_0}\int_{T_0} [r(t) - s(t; \psi)]^2 dt\right]. \tag{5.94}$$

Maximizing the likelihood function (5.94) results in the joint ML estimate of the carrier phase and symbol timing. Maximum likelihood estimation of the carrier phase and symbol timing can also be done separately, and

in subsequent sections we will discuss this separate estimation in more detail. Techniques for joint estimation are more complex; details of such techniques can be found in [11, Chap. 6.4] and [12, Chaps. 8–9].

5.6.2 Maximum Likelihood Phase Estimation

In this section we derive the maximum likelihood phase estimate assuming the timing is known. The likelihood function (5.94) with timing known reduces to

$$\begin{aligned}\Lambda(\phi) &= \exp\left[-\frac{1}{N_0}\int_{T_0}[r(t)-s(t;\phi)]^2\,dt\right] \\ &= \exp\left[-\frac{1}{N_0}\int_{T_0}r^2(t)\,dt+\frac{2}{N_0}\int_{T_0}r(t)s(t;\phi)\,dt-\frac{1}{N_0}\int_{T_0}s^2(t;\phi)\,dt\right].\end{aligned} \tag{5.95}$$

We estimate the carrier phase as the value $\hat{\phi}$ that maximizes this function. Note that the first term in (5.95) is independent of ϕ. Moreover, we assume that the third integral, which measures the energy in $s(t;\phi)$ over the observation interval, is relatively constant in ϕ. Given these assumptions, we see that the $\hat{\phi}$ that maximizes (5.95) also maximizes

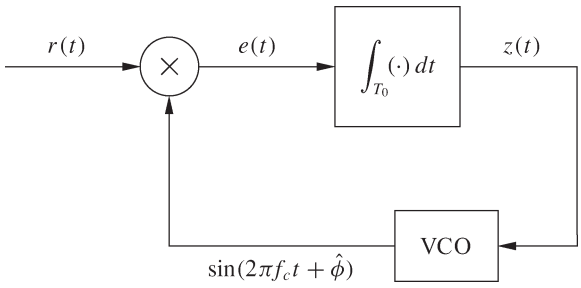


Figure 5.29: Phase-locked loop for carrier phase recovery (unmodulated carrier).

$$\Lambda'(\phi) = \int_{T_0} r(t)s(t;\phi)\,dt. \tag{5.96}$$

We can solve directly for the maximizing $\hat{\phi}$ in the simple case where the received signal is just an unmodulated carrier plus noise: $r(t) = A\cos(2\pi f_c t+\phi)+n(t)$. Then $\hat{\phi}$ must maximize

$$\Lambda'(\phi) = \int_{T_0} r(t)\cos(2\pi f_c t+\phi)\,dt. \tag{5.97}$$

Differentiating $\Lambda'(\phi)$ relative to ϕ and then setting it to zero yields that $\hat{\phi}$ satisfies

$$\int_{T_0} r(t)\sin(2\pi f_c t+\hat{\phi})\,dt = 0. \tag{5.98}$$

Solving (5.98) for $\hat{\phi}$ yields

$$\hat{\phi} = -\tan^{-1}\left[\frac{\int_{T_0} r(t)\sin(2\pi f_c t)\,dt}{\int_{T_0} r(t)\cos(2\pi f_c t)\,dt}\right]. \tag{5.99}$$

We could build a circuit to compute (5.99) from the received signal $r(t)$; in practice, however, carrier phase recovery is accomplished by using a phase lock loop to satisfy (5.98), as shown in Figure 5.29. In this figure, the integrator input in the absence of noise is given by $e(t) = r(t)\sin(2\pi f_c t + \hat{\phi})$ and the integrator output is

$$z(t) = \int_{T_0} r(t)\sin(2\pi f_c t + \hat{\phi})dt, \tag{5.100}$$

which is precisely the left-hand side of (5.98). Thus, if $z(t) = 0$ then the estimate $\hat{\phi}$ is the maximum likelihood estimate for ϕ. If $z(t) \neq 0$ then the voltage-controlled oscillator (VCO) adjusts its phase estimate $\hat{\phi}$ up or down depending on the polarity of $z(t)$: for $z(t) > 0$ it decreases $\hat{\phi}$ to reduce $z(t)$, and for $z(t) < 0$ it increases $\hat{\phi}$ to increase $z(t)$. In practice the integrator in Figure 5.29 is replaced with a *loop filter* whose output $.5A\sin(\hat{\phi} - \phi) \approx .5A(\hat{\phi} - \phi)$ is a function of the low-frequency component of its input $e(t) = A\cos(2\pi f_c t + \phi)\sin(2\pi f_c t + \hat{\phi}) = .5A\sin(\hat{\phi} - \phi) + .5A\sin(2\pi f_c t + \phi + \hat{\phi})$. This discussion of phase-locked loop (PLL) operation assumes that $\hat{\phi} \approx \phi$ because otherwise the polarity of $z(t)$ may not indicate the correct phase adjustment; that is, we would not necessarily have $\sin(\hat{\phi} - \phi) \approx \hat{\phi} - \phi$. The PLL typically exhibits some transient behavior in its initial estimation of the carrier phase. The advantage of a PLL is that it continually adjusts its estimate $\hat{\phi}$ to maintain $z(t) \approx 0$, which corrects for slow phase variations due to oscillator drift at the transmitter or changes in the propagation delay. In fact, the PLL is an example of a feedback control loop. More details on the PLL and its performance can be found in [11, 12].

The PLL derivation is for an unmodulated carrier, yet amplitude and phase modulation embed the message bits into the amplitude and phase of the carrier. For such signals there are two common carrier phase recovery approaches to deal with the effect of the data sequence on the received signal: the data sequence is either (a) assumed known or (b) treated as random with the phase estimate averaged over the data statistics. The first scenario is referred to as *decision-directed* parameter estimation, and this scenario typically results from sending a known training sequence. The second scenario is referred to as *non-decision directed* parameter estimation. With this technique the likelihood function (5.95) is maximized by averaging over the statistics of the data. One decision-directed technique uses data decisions to remove the modulation of the received signal: the resulting unmodulated carrier is then passed through a PLL. This basic structure is called a *decision-feedback PLL* because data decisions are fed back into the PLL for processing. The structure of a non–decision-directed carrier phase recovery loop depends on the underlying distribution of the data. For large constellations, most distributions lead to highly nonlinear functions of the parameter to be estimated. In this case the symbol distribution is often assumed to be Gaussian along each signal dimension, which greatly simplifies the recovery loop structure. An alternate non–decision-directed structure takes the Mth power of the signal ($M = 2$ for PAM and M for MPSK modulation), passes it through a bandpass filter at frequency Mf_c, and then uses a PLL. The nonlinear operation removes the effect of the amplitude or phase modulation so that the PLL can operate on an unmodulated carrier at frequency Mf_c. Many other structures for both decision-directed and non–decision-directed carrier recovery can be used, with different trade-offs in performance and complexity. A more comprehensive discussion of design and performance of carrier phase recovery can be found in [11, Chaps. 6.2.4–6.2.5] and [12].

5.6.3 Maximum Likelihood Timing Estimation

In this section we derive the maximum likelihood estimate of delay τ assuming the carrier phase is known. Since we assume that the phase ϕ is known, the timing recovery will not affect the carrier phase recovery loop and associated downconversion shown in Figure 5.28. Thus, it suffices to consider timing estimation for the in-phase or quadrature equivalent lowpass signals of $r(t)$ and $s(t;\tau)$. We denote the in-phase and quadrature components for $r(t)$ as $r_I(t)$ and $r_Q(t)$ and for $s(t;\tau)$ as $s_I(t;\tau)$ and $s_Q(t;\tau)$. We focus on the in-phase branch since the timing recovered from this branch can be used for the quadrature branch. The equivalent lowpass in-phase signal is given by

$$s_I(t;\tau) = \sum_k s_I(k) g(t - kT_s - \tau), \tag{5.101}$$

where $g(t)$ is the pulse shape and $s_I(k)$ denotes the amplitude associated with the in-phase component of the message transmitted over the kth symbol period. The in-phase equivalent lowpass received signal is $r_I(t) = s_I(t;\tau) + n_I(t)$. As in the case of phase synchronization, there are two categories of timing estimators: those for which the information symbols output from the demodulator are assumed known (decision-directed estimators), and those for which this sequence is not assumed known (non–decision-directed estimators).

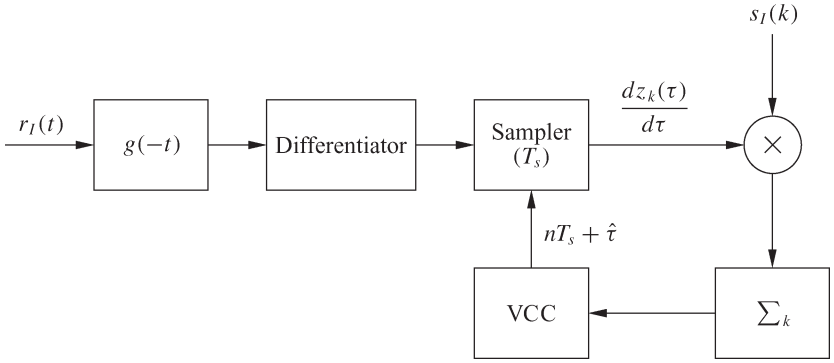


Figure 5.30: Decision-directed timing estimation.

The likelihood function (5.94) with known phase ϕ has a form similar to (5.95), the case of known delay:

$$\begin{aligned}\Lambda(\tau) &= \exp\left[-\frac{1}{N_0}\int_{T_0}[r_I(t) - s_I(t;\tau)]^2\,dt\right] \\ &= \exp\left[-\frac{1}{N_0}\int_{T_0} r_I^2(t)\,dt + \frac{2}{N_0}\int_{T_0} r_I(t)s_I(t;\tau)\,dt - \frac{1}{N_0}\int_{T_0} s_I^2(t;\tau)\,dt\right].\end{aligned} \tag{5.102}$$

Since the first and third terms in (5.102) do not change significantly with τ, the delay estimate $\hat{\tau}$ that maximizes (5.102) also maximizes

$$\begin{aligned}\Lambda'(\tau) &= \int_{T_0} r_I(t)s_I(t;\tau)\,dt \\ &= \sum_k s_I(k)\int_{T_0} r(t)g(t - kT_s - \tau)\,dt = \sum_k s_I(k)z_k(\tau),\end{aligned} \tag{5.103}$$

where

$$z_k(\tau) = \int_{T_0} r(t)g(t - kT_s - \tau)\,dt. \tag{5.104}$$

Differentiating (5.103) relative to τ and then setting it to zero yields that the timing estimate $\hat{\tau}$ must satisfy

$$\sum_k s_I(k)\frac{d}{d\tau}z_k(\tau) = 0. \tag{5.105}$$

For decision-directed estimation, (5.105) gives rise to the estimator shown in Figure 5.30. The input to the voltage-controlled clock (VCC) is (5.105). If this input is zero, then the timing estimate $\hat{\tau} = \tau$. If not, the clock (i.e., the timing estimate $\hat{\tau}$) is adjusted to drive the VCC input to zero. This timing estimation loop is also an example of a feedback control loop.

One structure for non–decision-directed timing estimation is the *early–late gate synchronizer* shown in Figure 5.31. This structure exploits two properties of the autocorrelation of $g(t)$, $R_g(\tau) = \int_0^{T_S} g(t)g(t-\tau)\,dt$ – namely, its symmetry ($R_g(\tau) = R_g(-\tau)$) and the fact that its maximum value is at $\tau = 0$. The input to the sampler in the upper branch of Figure 5.31 is proportional to the autocorrelation $R_g(\hat{\tau}-\tau+\delta) = \int_0^{T_S} g(t-\tau)g(t-\hat{\tau}-\delta)\,dt$, and the input to the sampler in the lower branch is proportional to the autocorrelation $R_g(\hat{\tau}-\tau-\delta) = \int_0^{T_S} g(t-\tau)g(t-\hat{\tau}+\delta)dt$. If $\hat{\tau} = \tau$ then, since $R_g(\delta) = R_g(-\delta)$, the input to the loop filter will be zero and the voltage-controlled clock will maintain its correct timing estimate. If $\hat{\tau} > \tau$ then $R_g(\hat{\tau} - \tau + \delta) < R_g(\hat{\tau} - \tau - \delta)$, and this negative input to the VCC will cause it to decrease its estimate of $\hat{\tau}$. Conversely, if $\hat{\tau} < \tau$ then $R_g(\hat{\tau} - \tau + \delta) > R_g(\hat{\tau} - \tau - \delta)$, and this positive input to the VCC will cause it to increase its estimate of $\hat{\tau}$.

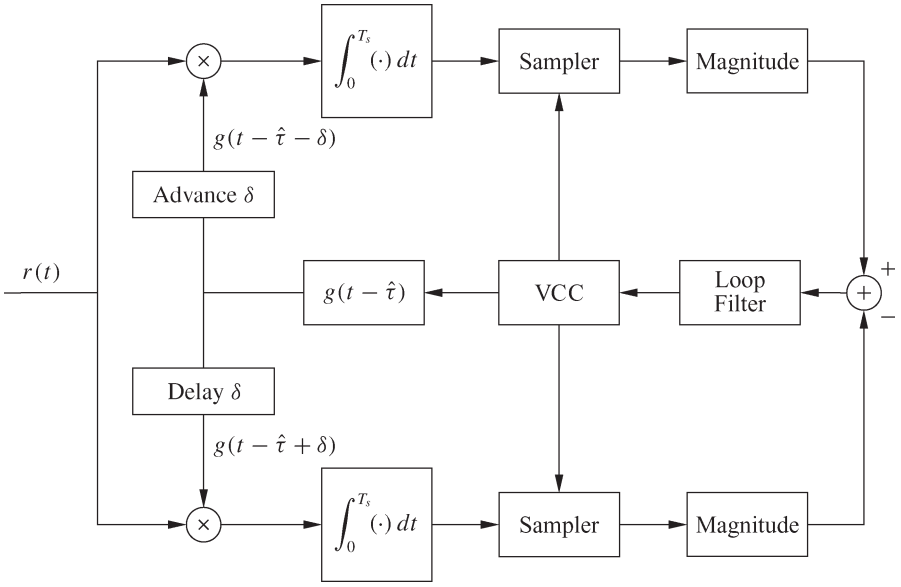


Figure 5.31: Early-late gate synchronizer.

More details on these and other structures for decision-directed and non–decision-directed timing estimation– as well as their performance trade-offs–can be found in [11, Chaps. 6.2.4–6.2.5] and [12].

Chapter 5 Problems

1. Using properties of orthonormal basis functions, show that if $s_i(t)$ and $s_j(t)$ have constellation points $\mathbf{s}_i$ and $\mathbf{s}_j$ (respectively) then

$$\|\mathbf{s}_i - \mathbf{s}_j\|^2 = \int_0^T (s_i(t) - s_j(t))^2 dt.$$

2. Find an alternate set of orthonormal basis functions for the space spanned by $\cos(2\pi t/T)$ and $\sin(2\pi t/T)$.

3. Consider a set of M orthogonal signal waveforms $s_m(t)$, for $1 \leq m \leq M$ and $0 \leq t \leq T$, where each waveform has the same energy $\mathcal{E}$. Define a new set of M waveforms as

$$s'_m(t) = s_m(t) - \frac{1}{M}\sum_{i=1}^{M} s_i(t), \quad 1 \leq m \leq M,\ 0 \leq t \leq T.$$

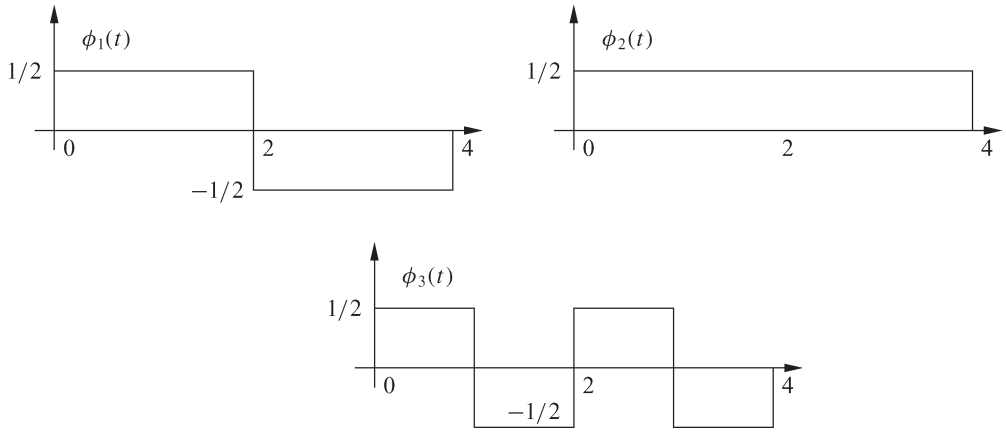


Figure 5.32: Signal waveforms for Problem 5-4.

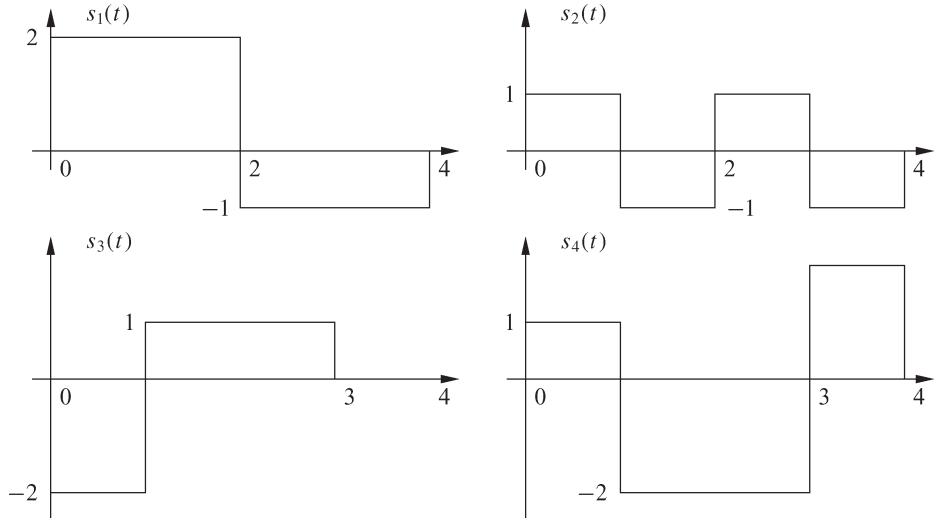


Figure 5.33: Signal waveforms for Problem 5-5.

Show that the M signal waveforms $\{s'_m(t)\}$ have equal energy, given by

$$\mathcal{E}' = (M-1)\mathcal{E}/M.$$

What is the inner product between any two waveforms?

4. Consider the three signal waveforms $\{\phi_1(t), \phi_2(t), \phi_3(t)\}$ shown in Figure 5.32.
   (a) Show that these waveforms are orthonormal.
   (b) Express the waveform $x(t)$ as a linear combination of $\{\phi_i(t)\}$ and find the coefficients, where $x(t)$ is given as

$$x(t) = \begin{cases} 2 & 0 \le t < 2, \\ 4 & 2 \le t \le 4. \end{cases}$$

5. Consider the four signal waveforms as shown in Figure 5.33.
   (a) Determine the dimensionality of the waveforms and a set of basis functions.
   (b) Use the basis functions to represent the four waveforms by vectors.
   (c) Determine the minimum distance between all the vector pairs.

6. Derive a mathematical expression for decision regions Z_i that minimize error probability assuming that messages are not equally likely – that is, assuming $p(m_i) = p_i (i = 1, \ldots, M)$, where p_i is not necessarily equal to $1/M$. Solve for these regions in the case of QPSK modulation with $s_1 = (A_c, 0)$, $s_2 = (0, A_c)$, $s_3 = (-A_c, 0)$, and $s_4 = (0, A_c)$, assuming $p(s_1) = p(s_3) = .2$ and $p(s_1) = p(s_3) = .3$.

7. Show that the remainder noise term $n_r(t)$ is independent of the correlator outputs r_i for all i. In other words, show that $\mathbf{E}[n_r(t) r_i] = 0$ for all i. Thus, since r_j (conditioned on $\mathbf{s}_i$) and $n_r(t)$ are Gaussian and uncorrelated, they are independent.

8. Show that output SNR is maximized when a given input signal is passed through a filter that is matched to that signal.

9. Find the matched filters $g(T-t), 0 \le t \le T$, and find $\int_0^T g(t) g(T-t)\, dt$ for the following waveforms.
   (a) Rectangular pulse: $g(t) = \sqrt{2/T}$.
   (b) Sinc pulse: $g(t) = \mathrm{sinc}(t)$.
   (c) Gaussian pulse: $g(t) = (\sqrt{\pi}/\alpha) e^{-\pi^2 t^2/\alpha^2}$.

10. Show that the ML receiver of Figure 5.4 is equivalent to the matched filter receiver of Figure 5.7.

11. Use the Chernoff Bound to derive the upper bound (5.46).

12. Compute the three bounds (5.43), (5.47), (5.48) as well as the approximation (5.49) for an asymmetric signal constellation $s_1 = (A_c, 0)$, $s_2 = (0, 2A_c)$, $s_3 = (-2A_c,\ 0)$, and $s_4 = (0, -A_c)$, assuming that $A_c/\sqrt{N_0} = 4$.

13. Find the input to each branch of the decision device in Figure 5.11 if the transmit carrier phase ϕ_0 differs from the receiver carrier phase ϕ by $\triangle\phi$.

14. Consider a 4-PSK constellation with $d_{\min} = \sqrt{2}$. What is the additional energy required to send one extra bit (8-PSK) while keeping the same minimum distance (and thus with the same bit error probability)?

15. Show that the average power of a square signal constellation with l bits per dimension, P_l, is proportional to $4^l/3$ and that the average power for one more bit per dimension, keeping the same minimum distance, is $P_{l+1} \approx 4P_l$. Find P_l for $l = 2$ and compute the average energy of MPSK and MPAM constellations with the same number of bits per symbol.

16. For MPSK with differential modulation, let $\triangle\phi$ denote the phase drift of the channel over a symbol time T_s. In the absence of noise, how large must $\triangle\phi$ be in order for a detection error to occur?

17. Find the Gray encoding of bit sequences to phase transitions in differential 8-PSK. Then find the sequence of symbols transmitted using differential 8-PSK modulation with this Gray encoding for the bit sequence 101110100101110 starting at the kth symbol time, assuming the transmitted symbol at the $(k-1)$th symbol time is $\mathbf{s}(k-1) = Ae^{j\pi/4}$.

18. Consider the octal signal point constellation shown in Figure 5.34.

    (a) The nearest neighbor signal points in the 8-QAM signal constellation are separated by a distance of A. Determine the radii a and b of the inner and outer circles.

    (b) The adjacent signal points in the 8-PSK are separated by a distance of A. Determine the radius r of the circle.

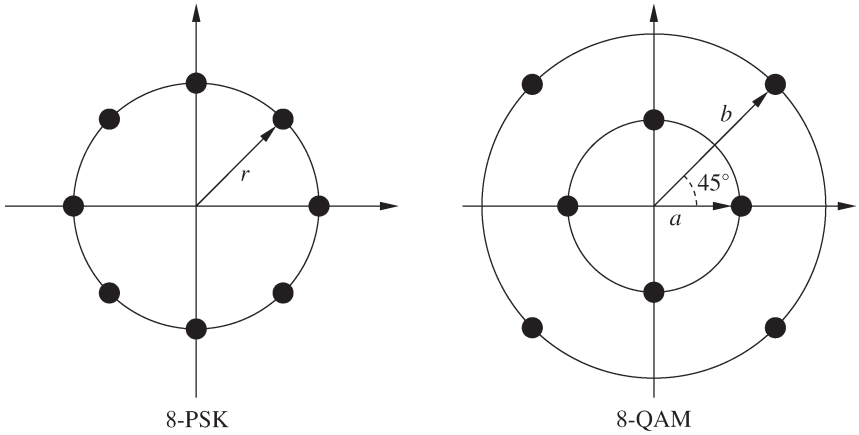


Figure 5.34: Octal signal point constellation for Problem 5-17.

    (c) Determine the average transmitter powers for the two signal constellations and compare the two powers. What is the relative power advantage of one constellation over the other? (Assume that all signal points are equally probable.)

    (d) Is it possible to assign three data bits to each point of the signal constellation such that nearest neighbor (adjacent) points differ in only one bit position?

    (e) Determine the symbol rate if the desired bit rate is 90 Mbps.

19. The $\pi/4$-QPSK modulation may be considered as two QPSK systems offset by $\pi/4$ radians.

    (a) Sketch the signal space diagram for a $\pi/4$-QPSK signal.

(b) Using Gray encoding, label the signal points with the corresponding data bits.

(c) Determine the sequence of symbols transmitted via $\pi/4$-QPSK for the bit sequence 0100100111100101.

(d) Repeat part (c) for $\pi/4$-DQPSK, assuming that the last symbol transmitted on the in-phase branch had a phase of π and that the last symbol transmitted on the quadrature branch had a phase of $-3\pi/4$.

20. Show that the minimum frequency separation for FSK such that the $\cos(2\pi f_j t)$ and $\cos(2\pi f_i t)$ are orthogonal is $\triangle f = \min_{ij} |f_j - f_i| = .5/T_s$.

21. Show that the Nyquist criterion for zero ISI pulses given by $p(kT_s) = p_0\delta(k)$ is equivalent to the frequency domain condition (5.79).

22. Show that the Gaussian pulse shape does not satisfy the Nyquist criterion.

Bibliography

[1] S. Haykin and M. Moher, *Communication Systems*, 5th ed., Wiley, New York, 2009.

[2] J. Proakis and M. Salehi, *Communication Systems Engineering*, 2nd ed., Pearson, 2001.

[3] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, New York, 1965.

[4] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics,* Vol. 23, pp. 493-507, 1952.

[5] M. Fitz, "Further results in the unified analysis of digital communication systems," *IEEE Trans. Commun*., pp. 521–32, March 1992.

[6] R. Ziemer, "An overview of modulation and coding for wireless communications," *Proc. IEEE Veh. Tech. Conf*., pp. 26–30, April 1996.

[7] G. D. Forney, Jr., and L.-F. Wei, "Multidimensional constellations-Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE J. Sel. Areas Commun*., pp. 877–92, August 1989.

[8] T. S. Rappaport, *Wireless Communications – Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

[9] G. L. Stuber, *Principles of Mobile Communications*, 4th ed., Springer, 2017.

[10] S. Haykin and M. Moher, *An Introduction to Analog and Digital Communications*, 2nd ed., Wiley, New York, 2006.

[11] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed., McGraw-Hill, New York, 2007.

[12] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum, New York, 1997.

[13] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers*, vol. 2, *Synchronization, Channel Estimation, and Signal Processing*, Wiley, New York, 1997.

[14] L. E. Franks, "Carrier and bit synchronization in data communication – A tutorial review," *IEEE Trans. Commun*., pp. 1107–21, August 1980.

Chapter 6

Performance of Digital Modulation over Wireless Channels

We now consider the performance of the digital modulation techniques discussed in the previous chapter when used over AWGN channels and channels with flat and frequency selective fading. For AWGN channels, the performance is dictated by the probability that a bit or symbol is received in error. This error probability is based on the modulation and the signal-to-noise power ratio (SNR). In flat-fading there are two performance criteria of interest depending on the channel coherance time; The average probability of error is used when the channel coherence time is on the order of the symbol time. The outage probability, defined as the probability that the instantaneous SNR falls below a given threshold, is used when the channel coherence time is much larger than the symbol time, so a deep fade affects many symbols. Wireless channels may also exhibit frequency-selective fading and Doppler frequency shift. Frequency-selective fading gives rise to intersymbol interference (ISI), which causes an irreducible error floor in the received signal. Doppler causes spectral broadening, which leads to adjacent channel interference (small at typical user velocities) and also to an irreducible error floor in signals with differential phase encoding (e.g. DPSK), since the phase reference of the previous symbol partially decorrelates over a symbol time. This chapter describes the impact on digital modulation performance of noise, flat fading, frequency-selective fading, and Doppler frequency shift.

6.1 AWGN Channels

In this section we define the signal-to-noise power ratio (SNR) and its relation to energy per bit (E_b) and energy per symbol (E_s). We then examine the error probability on AWGN channels for different modulation techniques as parameterized by these energy metrics. Our analysis uses the signal space concepts of Section 5.1.

6.1.1 Signal-to-Noise Power Ratio and Bit/Symbol Energy

In an AWGN channel the modulated signal $s(t) = \mathrm{Re}\{u(t)e^{j2\pi f_c t}\}$ has noise $n(t)$ added to it prior to reception. The noise $n(t)$ is a white Gaussian random process with mean zero and power spectral density (PSD) $N_0/2$. The received signal is thus $r(t) = s(t) + n(t)$.

We define the received SNR as the ratio of the received signal power P_r to the power of the noise within the bandwidth of the transmitted signal $s(t)$. The received power P_r is determined by the transmitted power and the path loss, shadowing, and multipath fading, as described in Chapters 2 and 3. The noise power is determined by the bandwidth of the transmitted signal and the spectral properties of $n(t)$. Specifically, if the bandwidth of the complex envelope $u(t)$ of $s(t)$ is B then the bandwidth of the transmitted signal $s(t)$ is $2B$. Since the noise $n(t)$

has uniform PSD $N_0/2$, the total noise power within the bandwidth $2B$ is $N = N_0/2 \cdot 2B = N_0 B$. Hence the received SNR is given by

$$\text{SNR} = \frac{P_r}{N_0 B}. \tag{6.1}$$

In systems with interference, we often use the received signal-to-interference-plus-noise power ratio (SINR) in place of SNR for calculating error probability. This is a reasonable approximation if the interference statistics approximate those of Gaussian noise. The received SINR is given by

$$\text{SINR} = \frac{P_r}{N_0 B + P_I}, \tag{6.2}$$

where P_I is the average power of the interference.

The SNR is often expressed in terms of the signal energy per bit E_b (or per symbol, E_s) as

$$\text{SNR} = \frac{P_r}{N_0 B} = \frac{E_s}{N_0 B T_s} = \frac{E_b}{N_0 B T_b}, \tag{6.3}$$

where T_s is the symbol time and T_b is the bit time (for binary modulation $T_s = T_b$ and $E_s = E_b$). Let R denote the modulation data rate in bits per second. We thus have that $E_b/N_0 = \text{SNR}/\eta$ where $\eta = R/B = 1/(BT_b)$ bps is called the *spectral efficiency* of the modulation. For pulse shaping with $T_s = 1/B$ (e.g., raised cosine pluses with $\beta = 1$), we have SNR $= E_s/N_0$ for multilevel signaling and SNR $= E_b/N_0$ for binary signaling. For general pulses, $T_s = k/B$ for some constant k, in which case $k \cdot \text{SNR} = E_s/N_0$.

The quantities $\gamma_s = E_s/N_0$ and $\gamma_b = E_b/N_0$ are sometimes called the SNR per symbol and the SNR per bit, respectively. For performance specification, we are interested in the bit error probability P_b as a function of γ_b. However, with M-ary signaling (e.g., MPAM and MPSK) the bit error probability depends on both the symbol error probability and the mapping of bits to symbols. Thus, we typically compute the symbol error probability P_s as a function of γ_s based on the signal space concepts of Section 5.1 and then obtain P_b as a function of γ_b using an exact or approximate conversion. The approximate conversion typically assumes that the symbol energy is divided equally among all bits and that Gray encoding is used, so that (at reasonable SNRs) one symbol error corresponds to exactly one bit error. These assumptions for M-ary signaling lead to the approximations

$$\gamma_b \approx \frac{\gamma_s}{\log_2 M} \tag{6.4}$$

and

$$P_b \approx \frac{P_s}{\log_2 M}. \tag{6.5}$$

6.1.2 Error Probability for BPSK and QPSK

We first consider BPSK modulation with coherent detection and perfect recovery of the carrier frequency and phase. With binary modulation each symbol corresponds to one bit, so the symbol and bit error rates are the same. The transmitted signal is $s_1(t) = Ag(t)\cos(2\pi f_c t)$ to send a 0-bit and $s_2(t) = -Ag(t)\cos(2\pi f_c t)$ to send a 1-bit for $A > 0$. From (5.54) we have that the probability of error is

$$P_b = Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right). \tag{6.6}$$

From Section 5.3.2, $d_{\min} = \|s_1 - s_0\| = \|A - (-A)\| = 2A$. Let us now relate A to the energy per bit. We have

$$E_b = \int_0^{T_b} s_1^2(t)\,dt = \int_0^{T_b} s_2^2(t)\,dt = \int_0^{T_b} A^2 g^2(t)\cos^2(2\pi f_c t)\,dt = A^2 \tag{6.7}$$

by (5.66). Thus, the signal constellation for BPSK in terms of energy per bit is given by $\mathbf{s}_0 = \sqrt{E_b}$ and $\mathbf{s}_1 = -\sqrt{E_b}$. This yields the minimum distance $d_{\min} = 2A = 2\sqrt{E_b}$. Substituting this into (6.6) yields

$$P_b = Q\left(\frac{2\sqrt{E_b}}{\sqrt{2N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\sqrt{2\gamma_b}\right), \tag{6.8}$$

which is plotted in Figure 6.1 below.

QPSK modulation consists of BPSK modulation on both the in-phase and quadrature components of the signal. With perfect phase and carrier recovery, the received signal components corresponding to each of these branches are orthogonal. Therefore, the bit error probability on each branch is the same as for BPSK: $P_b = Q(\sqrt{2\gamma_b})$. The symbol error probability equals the probability that either branch has a bit error:

$$P_s = 1 - \left[1 - Q\left(\sqrt{2\gamma_b}\right)\right]^2. \tag{6.9}$$

Since the symbol energy is split between the in-phase and quadrature branches, we have $\gamma_s = 2\gamma_b$. Substituting this into (6.9) yields P_s in terms of γ_s as

$$P_s = 1 - \left[1 - Q\left(\sqrt{\gamma_s}\right)\right]^2. \tag{6.10}$$

From Section 5.1.5, the union bound (5.43) on P_s for QPSK is

$$P_s \le 2Q(A/\sqrt{N_0}) + Q(\sqrt{2}A/\sqrt{N_0}). \tag{6.11}$$

Writing this in terms of $\gamma_s = 2\gamma_b = A^2/N_0$ yields

$$P_s \le 2Q(\sqrt{\gamma_s}) + Q(\sqrt{2\gamma_s}) \le 3Q(\sqrt{\gamma_s}). \tag{6.12}$$

The closed-form bound (5.48) becomes

$$P_s \le \frac{3}{\sqrt{2\pi\gamma_s}}\exp[-.5\gamma_s]. \tag{6.13}$$

Using the fact that the minimum distance between constellation points is $d_{\min} = \sqrt{2A^2}$ in (5.49), we obtain the nearest neighbor approximation

$$P_s \approx 2Q(\sqrt{A^2/N_0}) = 2Q(\sqrt{\gamma_s}). \tag{6.14}$$

Note that with Gray encoding we can approximate P_b from P_s by $P_b \approx P_s/2$, since QPSK has two bits per symbol.

Example 6.1: Find the bit error probability P_b and symbol error probability P_s of QPSK assuming $\gamma_b = 7$ dB. Compare the exact P_b with the approximation $P_b \approx P_s/2$ based on the assumption of Gray coding. Finally, compute P_s based on the nearest neighbor bound using $\gamma_s = 2\gamma_b$ and then compare with the exact P_s.

Solution: We have $\gamma_b = 10^{7/10} = 5.012$, so

$$P_b = Q(\sqrt{2\gamma_b}) = Q(\sqrt{10.024}) = 7.726 \cdot 10^{-4}. \tag{6.15}$$

The exact symbol error probability P_s is

$$P_s = 1 - [1 - Q(\sqrt{2\gamma_b})]^2 = 1 - [1 - Q(\sqrt{10.02})]^2 = 1.545 \cdot 10^{-3}. \tag{6.16}$$

The bit error probability approximation assuming Gray coding yields $P_b \approx P_s/2 = 7.723 \cdot 10^{-4}$, which is quite close to the exact P_b. The nearest neighbor approximation to P_s yields

$$P_s \approx 2Q(\sqrt{\gamma_s}) = 2Q(\sqrt{10.024}) = 1.545 \cdot 10^{-3}, \tag{6.17}$$

which matches well with the exact P_s.

6.1.3 Error Probability for MPSK

The signal constellation for MPSK has $s_{i1} = A\cos[2\pi(i-1)/M]$ and $s_{i2} = A\sin[2\pi(i-1)/M]$ for $A > 0$ and $i = 1, \ldots, M$. The symbol energy is $E_s = A^2$, so $\gamma_s = A^2/N_0$. From (5.67) it follows that, for the received vector $\mathbf{r} = re^{j\theta}$ represented in polar coordinates, an error occurs if the ith signal constellation point is transmitted and $\theta \notin (2\pi(i-1-.5)/M, 2\pi(i-1+.5)/M)$. The joint distribution of r and θ can be obtained through a bivariate transformation of the noise n_1 and n_2 on the in-phase and quadrature branches [1, Chap. 5.2.7], which yields

$$p(r,\theta) = \frac{r}{\pi N_0} \exp\left[-\frac{1}{N_0}(r^2 - 2\sqrt{E_s}r\cos(\theta) + E_s)\right]. \tag{6.18}$$

Since the error probability depends only on the distribution of θ, we can integrate out the dependence on r to obtain

$$p(\theta) = \int_0^\infty p(r,\theta)\, dr = \frac{1}{2\pi} e^{-\gamma_s \sin^2(\theta)} \int_0^\infty z \exp\left[-\frac{(z - \sqrt{2\gamma_s}\cos(\theta))^2}{2}\right] dz. \tag{6.19}$$

By symmetry, the probability of error is the same for each constellation point. Thus, we can derive P_s from the probability of error assuming the constellation point $\mathbf{s}_1 = (A,\ 0)$ is transmitted, which is

$$\begin{aligned} P_s &= 1 - \int_{-\pi/M}^{\pi/M} p(\theta) d\theta \\ &= 1 - \int_{-\pi/M}^{\pi/M} \frac{1}{2\pi} e^{-\gamma_s \sin^2(\theta)} \int_0^\infty z \exp\left[-\frac{(z - \sqrt{2\gamma_s}\cos(\theta))^2}{2}\right] dz. \end{aligned} \tag{6.20}$$

A closed-form solution to this integral does not exist for $M > 4$ and so the exact value of P_s must be computed numerically.

Each point in the MPSK constellation has two nearest neighbors at distance $d_{\min} = 2A\sin(\pi/M)$. Thus, the nearest neighbor approximation (5.49) to P_s is given by

$$P_s \approx 2Q(\sqrt{2}A\sin(\pi/M)/\sqrt{N_0}) = 2Q(\sqrt{2\gamma_s}\sin(\pi/M)). \tag{6.21}$$

To obtain P_b versus γ_b for MPSK we use the conversion in Equations 6.4 and 6.5. A plot of P_b versus γ_b for 8PSK and 16PSK using the approximation in Equation 6.21 along with P_b for BPSK given in Equation 6.8 is shown in Figure 6.1. We see from this figure that for a given γ_b, BPSK has the best performance and 16PSK has the worst performance. That is because the spacing between constellation points for a given γ_b decreases as the constellation size increases, leading to a higher P_b. This nearest neighbor approximation can differ significantly

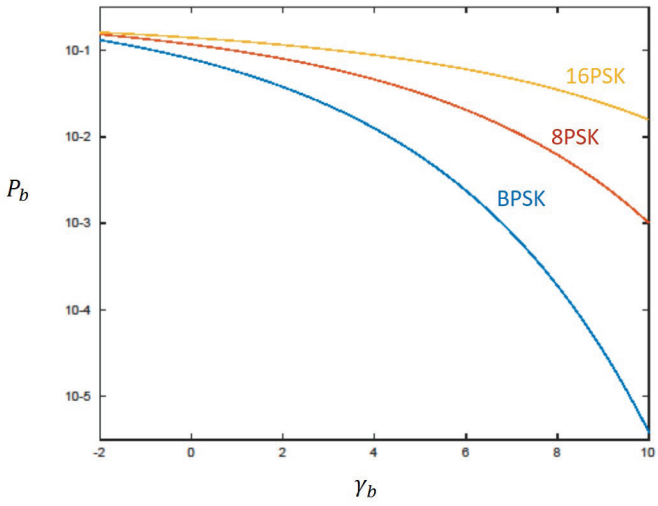


Figure 6.1: P_b versus γ_b for BPSK, 8PSK, and 16PSK in AWGN

from the exact value of P_s. However, it is much simpler to compute than the numerical integration of (6.20) that is required to obtain the exact P_s. This formula can also be obtained by approximating $p(\theta)$ as

$$p(\theta) \approx \sqrt{\gamma_s/\pi}\cos(\theta)e^{-\gamma_s \sin^2(\theta)}. \tag{6.22}$$

Using this in the first line of (6.20) yields (6.21).

Example 6.2: Compare the probability of bit error for 8-PSK and 16-PSK assuming $\gamma_b = 15$ dB and using the P_s approximation given in (6.21) along with the approximations (6.5) and (6.4).
Solution: From (6.4) we have that, for 8-PSK, $\gamma_s = (\log_2 8) \cdot 10^{15/10} = 94.87$. Substituting this into (6.21) yields

$$P_s \approx 2Q(\sqrt{189.74}\sin(\pi/8)) = 1.355 \cdot 10^{-7}. \tag{6.23}$$

Now, using (6.5), we get $P_b = P_s/3 = 4.52 \cdot 10^{-8}$. For 16-PSK we have $\gamma_s = (\log_2 16) \cdot 10^{15/10} = 126.49$. Substituting this into (6.21) yields

$$P_s \approx 2Q(\sqrt{252.98}\sin(\pi/16)) = 1.916 \cdot 10^{-3}, \tag{6.24}$$

and by using (6.5) we get $P_b = P_s/4 = 4.79 \cdot 10^{-4}$. Note that P_b is much larger for 16-PSK than for 8-PSK given the same γ_b. This result is expected because 16-PSK packs more bits per symbol into a given constellation and so, for a fixed energy per bit, the minimum distance between constellation points will be smaller.

The error probability derivation for MPSK assumes that the carrier phase is perfectly known to the receiver. Under phase estimation error, the distribution of $p(\theta)$ used to obtain P_s must incorporate the distribution of the phase rotation associated with carrier phase offset. This distribution is typically a function of the carrier phase estimation technique and the SNR. The impact of phase estimation error on coherent modulation is studied in [1, Apx. C], [2, Chap. 4.3.2], [3], and [4]. These works indicate that, as expected, significant phase offset leads to an irreducible bit error probability. Moreover, nonbinary signaling is more sensitive than BPSK to phase offset because of the resulting cross-coupling between in-phase and quadrature signal components. The impact of phase estimation

error can be especially severe in fast fading, where the channel phase changes rapidly owing to constructive and destructive multipath interference. Even with differential modulation, phase changes over and between symbol times can produce irreducible errors [5]. Timing errors can also degrade performance; analysis of timing errors in MPSK performance can be found in [2, Chap. 4.3.3] and [6].

6.1.4 Error Probability for MPAM and MQAM

The constellation for MPAM is $A_i = (2i - 1 - M)d$, $i = 1, 2, \ldots, M$. Each of the $M - 2$ inner constellation points of this constellation has two nearest neighbors at distance $2d$. The probability of making an error when sending one of these inner constellation points is just the probability that the noise exceeds d in either direction: $P_s(\mathbf{s}_i) = p(|\mathbf{n}| > d)$, $i = 2, \ldots, M - 1$. For the outer constellation points there is only one nearest neighbor, so an error occurs if the noise exceeds d in one direction only: $P_s(\mathbf{s}_i) = p(\mathbf{n} > d) = .5p(|\mathbf{n}| > d)$, $i = 1, M$. The probability of error is thus

$$\begin{aligned} P_s &= \frac{1}{M}\sum_{i=1}^{M} P_s(\mathbf{s}_i) \\ &= \frac{M-2}{M} 2Q\left(\sqrt{\frac{2d^2}{N_0}}\right) + \frac{2}{M}Q\left(\sqrt{\frac{2d^2}{N_0}}\right) \\ &= \frac{2(M-1)}{M}Q\left(\sqrt{\frac{2d^2}{N_0}}\right). \end{aligned} \tag{6.25}$$

From (5.62), the average energy per symbol for MPAM is

$$\overline{E}_s = \frac{1}{M}\sum_{i=1}^{M} A_i^2 = \frac{1}{M}\sum_{i=1}^{M}(2i - 1 - M)^2 d^2 = \frac{1}{3}(M^2 - 1)d^2. \tag{6.26}$$

Thus we can write P_s in terms of the average energy $\overline{E}_s$ as

$$P_s = \frac{2(M-1)}{M}Q\left(\sqrt{\frac{6\overline{\gamma}_s}{M^2-1}}\right). \tag{6.27}$$

Consider now MQAM modulation with a square signal constellation of size $M = L^2$. This system can be viewed as two MPAM systems with signal constellations of size L transmitted over the in-phase and quadrature signal components, each with half the energy of the original MQAM system. The constellation points in the in-phase and quadrature branches take values $A_i = (2i - 1 - L)d$, $i = 1, 2, \ldots, L$. The symbol error probability for each branch of the MQAM system is thus given by (6.27) with M replaced by $L = \sqrt{M}$ and $\overline{\gamma}_s$ equal to the average energy per symbol in the MQAM constellation:

$$P_{s,\,\text{branch}} = \frac{2(\sqrt{M}-1)}{\sqrt{M}}Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right). \tag{6.28}$$

Note that $\overline{\gamma}_s$ is multiplied by a factor of 3 in (6.28) instead of the factor of 6 in (6.27), since the MQAM constellation splits its total average energy $\overline{\gamma}_s$ between its in-phase and quadrature branches. The probability of symbol error for the MQAM system is then

$$P_s = 1 - \left(1 - \frac{2(\sqrt{M}-1)}{\sqrt{M}} Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right)\right)^2. \tag{6.29}$$

The nearest neighbor approximation to probability of symbol error depends on whether the constellation point is an inner or outer point. Inner points have four nearest neighbors, while outer points have either two or three nearest neighbors; in both cases the distance between nearest neighbors is $2d$. If we take a conservative approach and set the number of nearest neighbors to be four, we obtain the nearest neighbor approximation

$$P_s \approx 4Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right). \tag{6.30}$$

For nonrectangular constellations, it is relatively straightforward to show that the probability of symbol error is upper bounded as

$$P_s \leq 1 - \left[1 - 2Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right)\right]^2 \leq 4Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right), \tag{6.31}$$

which is the same as (6.30) for square constellations. The nearest neighbor approximation for nonrectangular constellations is

$$P_s \approx M_{d_{\min}} Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right), \tag{6.32}$$

where $M_{d_{\min}}$ is the largest number of nearest neighbors for any constellation point in the constellation and $d_{\min}$ is the minimum distance in the constellation.

Example 6.3: For 16-QAM with $\gamma_b = 15$ dB ($\gamma_s = \log_2 M \cdot \gamma_b$), compare the exact probability of symbol error (6.29) with (a) the nearest neighbor approximation (6.30) and (b) the symbol error probability for 16-PSK with the same γ_b (which was obtained in Example 6.2).

Solution: The average symbol energy $\gamma_s = 4 \cdot 10^{1.5} = 126.49$. The exact P_s is then given by

$$P_s = 1 - \left(1 - \frac{2(4-1)}{4} Q\left(\sqrt{\frac{3.126.49}{15}}\right)\right)^2 = 7.37 \cdot 10^{-7}. \tag{6.33}$$

The nearest neighbor approximation is given by

$$P_s \approx 4Q\left(\sqrt{\frac{3 \cdot 126.49}{15}}\right) = 9.82 \cdot 10^{-7}, \tag{6.34}$$

which is slightly larger than the exact value owing to the conservative approximation that every constellation point has four nearest neighbors. The symbol error probability for 16-PSK from Example 6.2 is $P_s \approx 1.916 \cdot 10^{-3}$, which is roughly four orders of magnitude larger than the exact P_s for 16-QAM. The larger P_s for MPSK versus MQAM with the same M and same γ_b is due to the fact that MQAM uses both amplitude and phase to encode data whereas MPSK uses just the phase. Thus, for the same energy per symbol or bit, MQAM makes more efficient use of energy and therefore has better performance.

---

3 The MQAM demodulator requires both amplitude and phase estimates of the channel so that the decision regions used in detection to estimate the transmitted symbol are not skewed in amplitude or phase. The analysis of performance degradation due to phase estimation error is similar to the case of MPSK discussed previously. The channel amplitude is used to scale the decision regions so that they correspond to the transmitted symbol: this scaling is called automatic gain control (AGC). If the channel gain is estimated in error then the AGC improperly scales the received signal, which can lead to incorrect demodulation even in the absence of noise. The channel gain is typically obtained using pilot symbols to estimate the channel gain at the receiver. However, pilot symbols do not lead to perfect channel estimates, and the estimation error can lead to bit errors. More details on the impact of amplitude and phase estimation errors on the performance of MQAM modulation can be found in [7, Chap. 10.3] and [8].

6.1.5 Error Probability for FSK and CPFSK

Let us first consider the error probability of binary FSK with the coherent demodulator of Figure 5.24. Since demodulation is coherent, we can neglect any phase offset in the carrier signals. The transmitted signal is defined by

$$s_i(t) = A\sqrt{2}T_b \cos(2\pi f_i t), \quad i = 1, 2. \tag{6.35}$$

Hence $E_b = A^2$ and $\gamma_b = A^2/N_0$. The input to the decision device is

$$z = s_1 + n_1 - s_2 - n_2. \tag{6.36}$$

The device outputs a 1-bit if $z > 0$ or a 0-bit if $z \leq 0$. Let us assume that $s_1(t)$ is transmitted; then

$$z|1 = A + n_1 - n_2. \tag{6.37}$$

An error occurs if $z = A + n_1 - n_2 \leq 0$. On the other hand, if $s_2(t)$ is transmitted then

$$z|0 = n_1 - A - n_2, \tag{6.38}$$

and an error occurs if $z = n_1 - A - n_2 > 0$. For n_1 and n_2 independent white Gaussian random variables with mean zero and variance $N_0/2$, their difference is a white Gaussian random variable with mean zero and variance equal to the sum of variances $N_0/2 + N_0/2 = N_0$. Then, for equally likely bit transmissions,

$$P_b = .5p(A + n_1 - n_2 \leq 0) + .5p(n_1 - A - n_2 > 0) = Q(A/\sqrt{N_0}) = Q(\sqrt{\gamma_b}). \tag{6.39}$$

The derivation of P_s for coherent MFSK with $M > 2$ is more complex and does not lead to a closed-form solution [2, eq. (4.92)]. The probability of symbol error for noncoherent MFSK is derived in [9, Chap. 8.1] as

$$P_s = \sum_{m=1}^{M-1} (-1)^{m+1} \binom{M-1}{m} \frac{1}{m+1} \exp\left[\frac{-m\gamma_s}{m+1}\right]. \tag{6.40}$$

The error probability of CPFSK depends on whether the detector is coherent or noncoherent and also on whether it uses symbol-by-symbol detection or sequence estimation. Analysis of error probability for CPFSK is complex because the memory in the modulation requires error probability analysis over multiple symbols. The formulas for error probability can also become quite complicated. Detailed derivations of error probability for these different CPFSK structures can be found in [1, Chap. 5.3]. As with linear modulations, FSK performance degrades under frequency and timing errors. A detailed analysis of the impact of such errors on FSK performance can be found in [2, Chap. 5.2], and [10, 11].

Table 6.1: Approximate symbol and bit error probabilities for coherent modulations

Modulation	$P_s(\gamma_s)$	$P_b(\gamma_b)$
BFSK		$P_b = Q(\sqrt{\gamma_b})$
BPSK		$P_b = Q(\sqrt{2\gamma_b})$
QPSK, 4-QAM	$P_s \approx 2Q(\sqrt{\gamma_s})$	$P_b = Q(\sqrt{2\gamma_b})$
MPAM	$P_s = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{6\overline{\gamma}_s}{M^2-1}}\right)$	$P_b \approx \frac{2(M-1)}{M \log_2 M} Q\left(\sqrt{\frac{6\overline{\gamma}_b \log_2 M}{M^2-1}}\right)$
MPSK	$P_s \approx 2Q\left(\sqrt{2\gamma_s} \sin\left(\frac{\pi}{M}\right)\right)$	$P_b \approx \frac{2}{\log_2 M} Q\left(\sqrt{2\gamma_b \log_2 M} \sin\left(\frac{\pi}{M}\right)\right)$
Rectangular MQAM	$P_s \approx 4Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right)$	$P_b \approx \frac{4}{\log_2 M} Q\left(\sqrt{\frac{3\overline{\gamma}_b \log_2 M}{M-1}}\right)$
Nonrectangular MQAM	$P_s \approx 4Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right)$	$P_b \approx \frac{4}{\log_2 M} Q\left(\sqrt{\frac{3\overline{\gamma}_b \log_2 M}{M-1}}\right)$

6.1.6 Error Probability Approximation for Coherent Modulations

Many of the approximations or exact values for P_s for coherent modulation are in the following form:

$$P_s(\gamma_s) \approx \alpha_M Q\left(\sqrt{\beta_M \gamma_s}\right), \tag{6.41}$$

where α_M and β_M depend on the type of approximation and the modulation type. In particular, the nearest neighbor approximation has this form, where α_M is the number of nearest neighbors to a constellation at the minimum distance and β_M is a constant that relates minimum distance to average symbol energy. In Table 6.1 we summarize the specific values of α_M and β_M for common P_s expressions for PSK, QAM, and FSK modulations based on the derivations in prior sections.

Performance specifications are generally more concerned with the bit error probability P_b as a function of the bit energy γ_b. To convert from P_s to P_b and from γ_s to γ_b we use the approximations (6.5) and (6.4), which assume Gray encoding and high SNR. Using these approximations in (6.41) yields a simple formula for P_b as a function of γ_b:

$$P_b(\gamma_b) = \hat{\alpha}_M Q\left(\sqrt{\hat{\beta}_M \gamma_b}\right), \tag{6.42}$$

where $\hat{\alpha}_M = \alpha_M / \log_2 M$ and $\hat{\beta}_M = (\log_2 M)\beta_M$ for α_M and β_M in (6.41). This conversion is used in what follows to obtain P_b versus γ_b from the general form of P_s versus γ_s in (6.41).

6.1.7 Error Probability for Differential Modulation

The probability of error for differential modulation is based on the phase difference associated with the phase comparator input of Figure 5.20. Specifically, the phase comparator extracts the phase of

$$\begin{aligned}\mathbf{r}(k)\mathbf{r}^*(k-1) =& A^2 e^{j(\theta(k)-\theta(k-1))} + A e^{j(\theta(k)+\phi_0)} n^*(k-1) \\ &+ A e^{-j(\theta(k-1)+\phi_0)} n(k) + n(k) n^*(k-1)\end{aligned} \tag{6.43}$$

in order to determine the transmitted symbol. By symmetry we can assume a given phase difference when computing the error probability. Assuming then a phase difference of zero, $\theta(k) - \theta(k-1) = 0$, yields

$$\mathbf{r}(k)\mathbf{r}^*(k-1) = A^2 + Ae^{j(\theta(k)+\phi_0)}n^*(k-1) + Ae^{-j(\theta(k-1)+\phi_0)}n(k) + n(k)n^*(k-1). \tag{6.44}$$

Next we define new random variables

$$\tilde{n}(k) = n(k)e^{-j(\theta(k-1)+\phi_0)} \quad \text{and} \quad \tilde{n}(k-1) = n(k-1)e^{-j(\theta(k)+\phi_0)}, \tag{6.45}$$

which have the same statistics as $n(k)$ and $n(k-1)$. Then

$$\mathbf{r}(k)\mathbf{r}^*(k-1) = A^2 + A(\tilde{n}^*(k-1) + \tilde{n}(k)) + \tilde{n}(k)\tilde{n}^*(k-1). \tag{6.46}$$

There are three terms in (6.46): the first term, with the desired phase difference of zero; and the second and third terms, which contribute noise. At reasonable SNRs the third noise term is much smaller than the second, so we neglect it. Dividing the remaining terms by A yields

$$\tilde{z} = A + \mathrm{Re}\{\tilde{n}^*(k-1) + \tilde{n}(k)\} + j\ \mathrm{Im}\{\tilde{n}^*(k-1) + \tilde{n}(k)\}. \tag{6.47}$$

Let us define $x = \mathrm{Re}\{\tilde{z}\}$ and $y = \mathrm{Im}\{\tilde{z}\}$. The phase of $\tilde{z}$ is then given by

$$\theta_{\tilde{z}} = \tan^{-1} y/x. \tag{6.48}$$

Given that the phase difference was zero, an error occurs if $|\theta_{\tilde{z}}| \geq \pi/M$. Determining $p(|\theta_{\tilde{z}}| \geq \pi/M)$ is identical to the case of coherent PSK except that, by (6.47), we have two noise terms instead of one and so the noise power is twice that of the coherent case. This will lead to a performance of differential modulation that is roughly 3 dB worse than that of coherent modulation.

In DPSK modulation we need only consider the in-phase branch of Figure 5.20 when making a decision, so we set $x = \mathrm{Re}\{\tilde{z}\}$ in our analysis. In particular, assuming a zero is transmitted, if $x = A + \mathrm{Re}\{\tilde{n}^*(k-1) + \tilde{n}(k)\} < 0$ then a decision error is made. This probability can be obtained by finding the characteristic or moment generating function for x, taking the inverse Laplace transform to get the distribution of x, and then integrating over the decision region $x < 0$. This technique is quite general and can be applied to a wide variety of different modulation and detection types in both AWGN and fading [9, Chap. 1.1]: we will use it later to compute the average probability of symbol error for linear modulations in fading both with and without diversity. In DPSK the characteristic function for x is obtained using the general quadratic form of complex Gaussian random variables (see [1, Apx. B] and [12, Apx. B]), and the resulting bit error probability is given by

$$P_b = \frac{1}{2}e^{-\gamma_b}. \tag{6.49}$$

The error probability of DQPSK is harder to obtain than that of DPSK, since the phase difference between two adjacent symbols depends on two information bits instead of one. An elegant analysis of the resulting error probability is obtained in [1, Apx. B] as a special case of the probability analysis for quadratic functions of complex Gaussian random variables. This yields, under Gray coding, the bit error probability [1, Eqn. 4.5-66]

$$P_b = Q_1(a,b) - \frac{1}{2}\exp\left[\frac{-(a^2+b^2)}{2}\right] I_0(ab), \tag{6.50}$$

where $Q_1(a,b)$ is the *Marcum Q function* defined as

$$Q_1(a,b) \triangleq \int_b^\infty z\exp\left[\frac{-(a^2+z^2)}{2}\right] I_0(az)dz, \tag{6.51}$$

for $I_0(x)$ the modified Bessel function of the first kind and zeroth order, $a = \sqrt{2\gamma_b(1-\sqrt{.5})}$, and $b = \sqrt{2\gamma_b(1+\sqrt{.5})}$.

6.2 Alternate Q-Function Representation

In (6.41) we saw that P_s for many coherent modulation techniques in AWGN is approximated in terms of the Gaussian Q-function. Recall that $Q(z)$ is defined as the probability that a Gaussian random variable X with mean 0 and variance 1 exceeds the value z:

$$Q(z) = p(X \geq z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \tag{6.52}$$

The Q-function is not that easy to work with since the argument z is in the lower limit of the integrand, the integrand has infinite range, and the exponential function in the integral doesn't lead to a closed-form solution.

In 1991 an alternate representation of the Q-function was obtained by Craig [13]. The alternate form is given by

$$Q(z) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left[\frac{-z^2}{2\sin^2\phi}\right] d\phi, \quad z > 0. \tag{6.53}$$

This representation can also be deduced from the work of Weinstein [14] or Pawula et al. [5]. In this alternate form, the integrand is over a finite range that is independent of the function argument z, and the integral is Gaussian with respect to z. These features will prove important in using the alternate representation to derive average error probability in fading.

Craig's motivation for deriving the alternate representation was to simplify the probability of error calculation for AWGN channels. In particular, we can write the probability of bit error for BPSK using the alternate form as

$$P_b = Q\left(\sqrt{2\gamma_b}\right) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left[\frac{-\gamma_b}{\sin^2\phi}\right] d\phi. \tag{6.54}$$

Similarly, the alternate representation can be used to obtain a simple *exact* formula for the P_s of MPSK in AWGN as

$$P_s = \frac{1}{\pi} \int_0^{(M-1)\pi/M} \exp\left[\frac{-\beta_M \gamma_s}{\sin^2\phi}\right] d\phi \tag{6.55}$$

(see [13]), where $\beta_M = \sin^2(\pi/M)$. Note that this formula does not correspond to the general form $\alpha_M Q(\sqrt{\beta_M \gamma_s})$: the general form is an approximation, whereas (6.55) is exact. Note also that (6.55) is obtained via a finite-range integral of simple trigonometric functions that is easily computed using a numerical computer package or calculator.

6.3 Fading

In AWGN the probability of symbol error depends on the received SNR or, equivalently, on γ_s. In a fading environment the received signal power varies randomly over distance or time as a result of shadowing and/or multipath fading. Thus, in fading, γ_s is a random variable with distribution $p_{\gamma_s}(\gamma)$ and so $P_s(\gamma_s)$ is also random. The performance metric when γ_s is random depends on the rate of change of the fading. There are three different performance criteria that can be used to characterize the random variable P_s:

- the outage probability, P_{out}, defined as the probability that γ_s falls below a given value γ_0 corresponding to the maximum allowable P_s;
- the average error probability, $\bar{P}_s$, averaged over the distribution of γ_s;

- combined average error probability and outage, defined as the average error probability that can be achieved some percentage of time or some percentage of spatial locations.

The average probability of symbol error applies when the fading coherence time is on the order of a symbol time ($T_s \approx T_c$), so that the signal fade level is roughly constant over a symbol period. Since many error correction coding techniques can recover from a few bit errors and since end-to-end performance is typically not seriously degraded by a few simultaneous bit errors (since the erroneous bits can be dropped or retransmitted), the average error probability is a reasonably good figure of merit for the channel quality under these conditions.

However, if the signal fading is changing slowly ($T_s \ll T_c$) then a deep fade will affect many simultaneous symbols. Hence fading may lead to large error bursts, which cannot be corrected with coding of reasonable complexity. Therefore, these error bursts can seriously degrade end-to-end performance. In this case acceptable performance cannot be guaranteed over all time – or, equivalently, throughout a cell – without drastically increasing transmit power. Under these circumstances, an outage probability is specified so that the channel is deemed unusable for some fraction of time or space. Outage and average error probability are often combined when the channel is modeled as a combination of fast and slow fading (e.g., log-normal shadowing with fast Rayleigh fading).

Note that if $T_c \ll T_s$ then the fading will be averaged out by the matched filter in the demodulator. Thus, for very fast fading, performance is the same as in AWGN.

6.3.1 Outage Probability

The outage probability relative to γ_0 is defined as

$$P_{\text{out}} = p(\gamma_s < \gamma_0) = \int_0^{\gamma_0} p_{\gamma_s}(\gamma) d\gamma, \tag{6.56}$$

where γ_0 typically specifies the minimum SNR required for acceptable performance. For example, if we consider digitized voice, $P_b = 10^{-3}$ is an acceptable error rate because it generally can't be detected by the human ear. Thus, for a BPSK signal in Rayleigh fading, $\gamma_b < 7\,\text{dB}$ would be declared an outage; hence we set $\gamma_0 = 7\,\text{dB}$.

In Rayleigh fading the outage probability becomes

$$P_{\text{out}} = \int_0^{\gamma_0} \frac{1}{\bar{\gamma}_s} e^{-\gamma_s/\bar{\gamma}_s} d\gamma_s = 1 - e^{-\gamma_0/\bar{\gamma}_s}. \tag{6.57}$$

Inverting this formula shows that, for a given outage probability, the required average SNR $\bar{\gamma}_s$ is

$$\bar{\gamma}_s = \frac{\gamma_0}{-\ln(1 - P_{\text{out}})}. \tag{6.58}$$

In decibels this means that $10 \log \bar{\gamma}_s$ must exceed the target $10 \log \gamma_0$ by

$$F_d = -10 \log[-\ln(1 - P_{\text{out}})] \tag{6.59}$$

in order to maintain acceptable performance more than $100 \cdot (1 - P_{\text{out}})$ percent of the time. The quantity F_d is typically called the *dB fade margin*.

Example 6.4: Determine the required $\bar{\gamma}_b$ for BPSK modulation in slow Rayleigh fading such that, for 95% of the time (or in 95% of the locations), $P_b(\gamma_b) < 10^{-4}$.

Solution: For BPSK modulation in AWGN the target BER is obtained at $\gamma_b = 8.5\,\text{dB}$. That is, for $P_b(\gamma_b) = Q(\sqrt{2\gamma_b})$ we have $P_b(10^{.85}) = 10^{-4}$. Thus, $\gamma_0 = 8.5\,\text{dB}$. We want $P_{\text{out}} = p(\gamma_b < \gamma_0) = .05$, so

$$\bar{\gamma}_b = \frac{\gamma_0}{-\ln(1 - P_{\text{out}})} = \frac{10^{.85}}{-\ln(1 - .05)} = 21.4\,\text{dB}. \tag{6.60}$$

6.3.2 Average Probability of Error

The average probability of error is used as a performance metric when $T_s \approx T_c$. We can therefore assume that γ_s is roughly constant over a symbol time. Then the average probability of error is computed by integrating the error probability in AWGN over the fading distribution:

$$\bar{P}_s = \int_0^\infty P_s(\gamma) p_{\gamma_s}(\gamma)\, d\gamma, \tag{6.61}$$

where $P_s(\gamma)$ is the probability of symbol error in AWGN with SNR γ, which can be approximated by the expressions in Table 6.1. For a given distribution of the fading amplitude r (e.g., Rayleigh, Rician, log-normal), we compute $p_{\gamma_s}(\gamma)$ by making the change of variable

$$p_{\gamma_s}(\gamma) d\gamma = p(r)\, dr. \tag{6.62}$$

For example, in Rayleigh fading the received signal amplitude r has the Rayleigh distribution

$$p(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0, \tag{6.63}$$

and the signal power is exponentially distributed with mean $2\sigma^2$. The SNR per symbol for a given amplitude r is

$$\gamma = \frac{r^2 T_s}{2\sigma_n^2}, \tag{6.64}$$

where $\sigma_n^2 = N_0/2$ is the PSD of the noise in the in-phase and quadrature branches. Differentiating both sides of this expression yields

$$d\gamma = \frac{r T_s}{\sigma_n^2}\, dr. \tag{6.65}$$

Substituting (6.64) and (6.65) into (6.63) and then (6.62) yields

$$p_{\gamma_s}(\gamma) = \frac{\sigma_n^2}{\sigma^2 T_s} e^{-\gamma \sigma_n^2/\sigma^2 T_s}. \tag{6.66}$$

Since the average SNR per symbol $\bar{\gamma}_s$ is just $\sigma^2 T_s/\sigma_n^2$, we can rewrite (6.66) as

$$p_\gamma(\gamma) = \frac{1}{\bar{\gamma}_s} e^{-\gamma/\bar{\gamma}_s}, \tag{6.67}$$

which is the exponential distribution. For binary signaling this reduces to

$$p_{\gamma_b}(\gamma) = \frac{1}{\bar{\gamma}_b} e^{-\gamma/\bar{\gamma}_b}. \tag{6.68}$$

Integrating (6.8) over the distribution (6.68) yields the following average probability of error for BPSK in Rayleigh fading:

$$\bar{P}_b = \frac{1}{2}\left[1 - \sqrt{\frac{\bar{\gamma}_b}{1+\bar{\gamma}_b}}\right] \approx \frac{1}{4\bar{\gamma}_b} \quad \text{(BPSK)}, \tag{6.69}$$

where the approximation holds for large $\bar{\gamma}_b$. A similar integration of (6.39) over (6.68) yields the average probability of error for binary FSK in Rayleigh fading as

$$\bar{P}_b = \frac{1}{2}\left[1 - \sqrt{\frac{\bar{\gamma}_b}{2+\bar{\gamma}_b}}\right] \approx \frac{1}{2\bar{\gamma}_b} \quad \text{(binary FSK)}. \tag{6.70}$$

For noncoherent modulation, if we assume the channel phase is relatively constant over a symbol time then we obtain the probability of error by again integrating the error probability in AWGN over the fading distribution. For DPSK this yields

$$\bar{P}_b = \frac{1}{2(1+\bar{\gamma}_b)} \approx \frac{1}{2\bar{\gamma}_b} \text{ (DPSK)}, \tag{6.71}$$

where again the approximation holds for large $\bar{\gamma}_b$. Thus, at high SNRs, the performances of DPSK and binary FSK converge, and this performance is approximately 3 dB worse than that of BPSK. This was also observed in AWGN and represents the power penalty of differential detection. In practice the power penalty is somewhat smaller, since DPSK can correct for slow phase changes introduced in the channel or receiver, which are not taken into account in these error calculations.

If we use the general approximation $P_s \approx \alpha_M Q(\sqrt{\beta_M \gamma_s})$ then the average probability of symbol error in Rayleigh fading can be approximated as

$$\bar{P}_s \approx \int_0^\infty \alpha_M Q(\sqrt{\beta_M \gamma}) \cdot \frac{1}{\bar{\gamma}_s} e^{-\gamma/\bar{\gamma}_s}\, d\gamma_s = \frac{\alpha_M}{2}\left[1 - \sqrt{\frac{.5\beta_M \bar{\gamma}_s}{1+.5\beta_M \bar{\gamma}_s}}\right] \approx \frac{\alpha_M}{2\beta_M \bar{\gamma}_s}, \tag{6.72}$$

where the last approximation is in the limit of high SNR.

It is interesting to compare the bit error probabilities of the different modulation schemes in AWGN and in fading. For binary PSK, FSK, and DPSK, the bit error probability in AWGN decreases exponentially with increasing γ_b. However, in fading the bit error probability for all the modulation types decreases just linearly with increasing $\bar{\gamma}_b$. Similar behavior occurs for nonbinary modulation. Thus, the power necessary to maintain a given P_b, particularly for small values, is much higher in fading channels than in AWGN channels. For example, in Figure 6.2 we plot the error probability of BPSK in AWGN and in flat Rayleigh fading. We see that it requires approximately 8-dB SNR to maintain a 10^{-3} bit error rate in AWGN, whereas it takes approximately 24-dB SNR to maintain the same error rate in fading. A similar plot for the error probabilities of MQAM, based on the approximations (6.30) and (6.72), is shown in Figure 6.3. From these figures it is clear that minimizing transmit power requires some technique to remove the effects of fading. We will discuss some of these techniques – including diversity combining, spread spectrum, and RAKE receivers – in later chapters.

Rayleigh fading is one of the worst-case fading scenarios. In Figure 6.4 we show the average bit error probability of BPSK in Nakagami fading for different values of the Nakagami-m parameter. We see that, as m increases, the fading decreases and the average bit error probability converges to that of an AWGN channel.

6.3.3 Moment Generating Function Approach to Average Error Probability

The *moment generating function* (MGF) is a useful tool for performance analysis of modulation in fading both with and without diversity. In this section we discuss how it can be used to simplify performance analysis of average

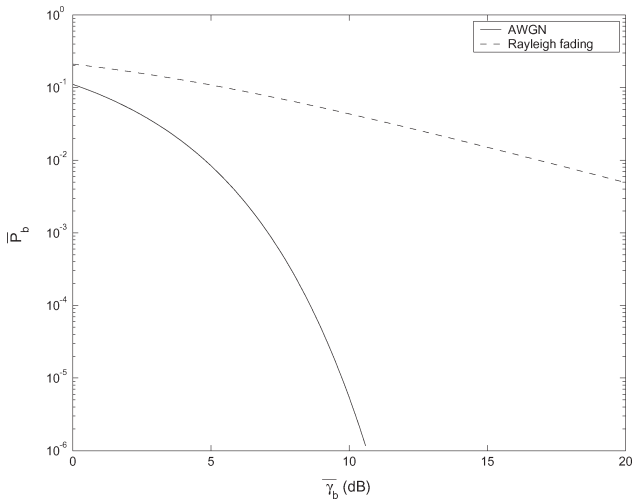


Figure 6.2: Average P_b for BPSK in Rayleigh fading and AWGN.

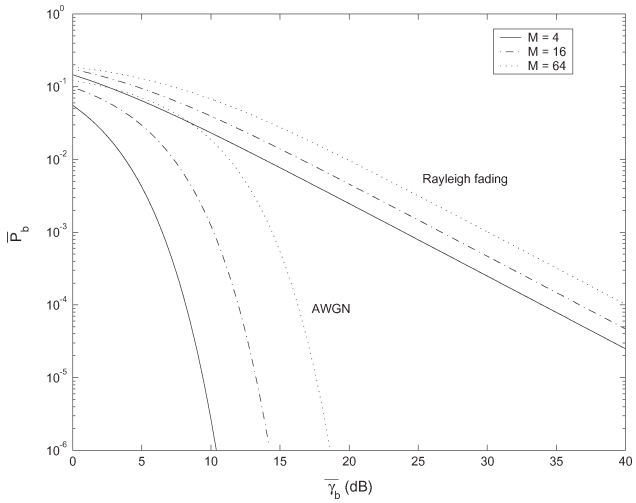


Figure 6.3: Average P_b for MQAM in Rayleigh fading and AWGN.

probability of symbol error in fading. In the next chapter we will see that it also greatly simplifies analysis in fading channels with diversity.

The MGF for a nonnegative random variable γ with distribution $p_\gamma(\gamma), \gamma \geq 0$, is defined as

$$\mathcal{M}_\gamma(s) = \int_0^\infty p_\gamma(\gamma) e^{s\gamma}\, d\gamma. \tag{6.73}$$

Note that this function is just the Laplace transform of the distribution $p_\gamma(\gamma)$ with the argument reversed in sign: $\mathcal{L}[p_\gamma(\gamma)] = \mathcal{M}_\gamma(-s)$. Thus, the MGF for most fading distributions of interest can be computed either in closed-form using classical Laplace transforms or through numerical integration. In particular, the MGF for common multipath fading distributions are as follows [9, Chap. 5.1].

■ Rayleigh:

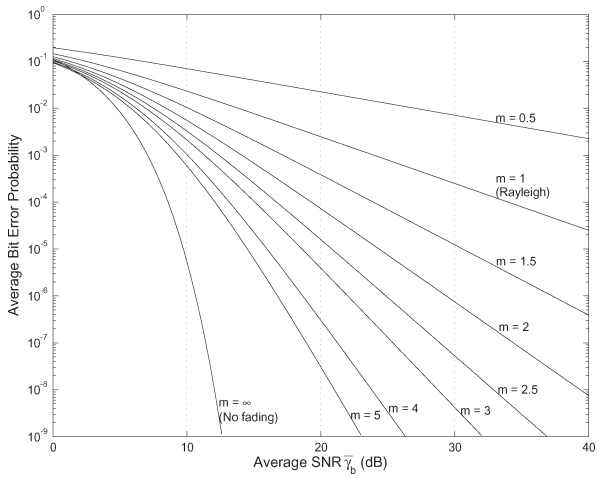


Figure 6.4: Average P_b for BPSK in Nakagami fading.

$$\mathcal{M}_{\gamma_s}(s) = (1 - s\bar{\gamma}_s)^{-1}. \tag{6.74}$$

■ Rician with factor K:

$$\mathcal{M}_{\gamma_s}(s) = \frac{1+K}{1+K-s\bar{\gamma}_s} \exp\left[\frac{Ks\bar{\gamma}_s}{1+K-s\bar{\gamma}_s}\right]. \tag{6.75}$$

■ Nakagami-m:

$$\mathcal{M}_{\gamma_s}(s) = \left(1 - \frac{s\bar{\gamma}_s}{m}\right)^{-m}. \tag{6.76}$$

As indicated by its name, the moments $\mathbf{E}[\gamma^n]$ of γ can be obtained from $\mathcal{M}_\gamma(s)$ as

$$\mathbf{E}[\gamma^n] = \frac{\partial^n}{\partial s^n}[\mathcal{M}_{\gamma_s}(s)]|_{s=0}. \tag{6.77}$$

The basic premise of the MGF approach for computing average error probability in fading is to express the probability of error P_s in AWGN for the modulation of interest either as an exponential function of γ_s,

$$P_s = a\exp[-b\gamma_s] \tag{6.78}$$

for constants a and b, or as a finite-range integral of such an exponential function:

$$P_s = \int_0^B a\exp[-b(x)\gamma_s]\,dx, \tag{6.79}$$

where the constant $b(x)$ may depend on the integrand but the SNR γ_s does not (and is not in the limits of integration). These forms allow the average probability of error to be expressed in terms of the MGF for the fading distribution. Specifically, if $P_s = a\exp[-b\gamma_s]$, then

$$\bar{P}_s = \int_0^\infty a\exp[-b\gamma]p_{\gamma_s}(\gamma)d\gamma = a\mathcal{M}_{\gamma_s}(-b). \tag{6.80}$$

Since DPSK is in this form with $a = 1/2$ and $b = 1$, we see that the average probability of bit error for DPSK in any type of fading is

$$\bar{P}_b = \tfrac{1}{2}\mathcal{M}_{\gamma_s}(-1), \tag{6.81}$$

where $\mathcal{M}_{\gamma_s}(s)$ is the MGF of the fading distribution. For example, using $\mathcal{M}_{\gamma_s}(s)$ for Rayleigh fading given by (6.74) with $s = -1$ yields $\bar{P}_b = [2(1 + \bar{\gamma}_b)]^{-1}$, which is the same as we obtained in (6.71). If P_s is in the integral form of (6.79) then

$$\begin{aligned}\bar{P}_s &= \int_0^\infty \int_0^B a \exp[-b(x)\gamma] dx \; p_{\gamma_s}(\gamma)\, d\gamma \\ &= a \int_0^B \left[\int_0^\infty \exp[-b(x)\gamma] p_{\gamma s}(\gamma) d\gamma \right] dx \\ &= a \int_0^B \mathcal{M}_{\gamma_s}(-b(x))\, dx. \end{aligned} \tag{6.82}$$

In this latter case, the average probability of symbol error is a single finite-range integral of the MGF of the fading distribution, which typically can be found in closed form or easily evaluated numerically.

Let us now apply the MGF approach to specific modulations and fading distributions. Consider any modulation for which its symbol error probability P_s in AWGN is given or approximated by an expression of the following form:

$$P_s(\gamma_s) = \alpha Q(\sqrt{\beta \gamma_s}). \tag{6.83}$$

Note that (6.41) provided an approximation to P_s in AWGN for coherent MPSK, MQAM, and MPAM in this form, with constants α and β that depend on M.

Using the alternate Q-function representation (6.53), we get that

$$P_s = \frac{\alpha}{\pi} \int_0^{\pi/2} \exp\left[\frac{-.5\beta\gamma}{\sin^2 \phi} \right] d\phi, \tag{6.84}$$

which is in the desired form (6.79). Thus, the average error probability in fading for modulations with $P_s = \alpha Q(\sqrt{\beta\gamma_s})$ in AWGN is given by

$$\begin{aligned}\bar{P}_s &= \frac{\alpha}{\pi} \int_0^\infty \int_0^{\pi/2} \exp\left[\frac{-.5\beta\gamma}{\sin^2 \phi} \right] p_{\gamma_s}(\gamma)\, d\gamma d\phi \\ &= \frac{\alpha}{\pi} \int_0^{\pi/2} \left[\int_0^\infty \exp\left[\frac{-.5\beta\gamma}{\sin^2 \phi} \right] p_{\gamma_s}(\gamma) d\gamma \right] d\phi \\ &= \frac{\alpha}{\pi} \int_0^{\pi/2} \mathcal{M}_{\gamma_s}\left(\frac{-.5\beta}{\sin^2 \phi} \right) d\phi, \end{aligned} \tag{6.85}$$

where $\mathcal{M}_{\gamma_s}(s)$ is the MGF associated with the distribution $p_{\gamma_s}(\gamma)$ as defined by (6.73). Recall that Table 6.1 approximates the error probability in AWGN for many modulations of interest as $P_s \approx \alpha Q(\sqrt{\beta\gamma_s})$, so (6.85) gives an approximation for the average error probability of these modulations in fading. Moreover, the exact average probability of symbol error for coherent MPSK can be obtained in a form similar to (6.85) by noting that Craig's formula for P_s of MPSK in AWGN given by (6.55) is in the desired form (6.79). Thus, the exact average probability of error for MPSK becomes

$$\begin{aligned}\bar{P}_s &= \int_0^{\infty} \frac{1}{\pi} \int_0^{(M-1)\pi/M} \exp\left[\frac{-\beta_M \gamma_s}{\sin^2\phi}\right] d\phi \, p_{\gamma_s}(\gamma)\, d\gamma \\ &= \frac{1}{\pi} \int_0^{(M-1)\pi/M} \left[\int_0^{\infty} \exp\left[\frac{-\beta_M \gamma_s}{\sin^2\phi}\right] p_{\gamma_s}(\gamma)\, d\gamma\right] d\phi \\ &= \frac{1}{\pi} \int_0^{(M-1)\pi/M} \mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) d\phi, \end{aligned} \tag{6.86}$$

where $\beta_M = \sin^2(\pi/M)$ depends on the size of the MPSK constellation. The MGF $\mathcal{M}_{\gamma_s}(s)$ for Rayleigh, Rician, and Nakagami-m distributions were given by (6.74), (6.75), and (6.76), respectively. Substituting $s = -\beta_M/\sin^2\phi$ in these expressions yields the following equations.

■ Rayleigh:

$$\mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) = \left(1 + \frac{\beta_M \bar{\gamma}_s}{\sin^2\phi}\right)^{-1}. \tag{6.87}$$

■ Rician with factor K:

$$\mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) = \frac{(1+K)\sin^2\phi}{(1+K)\sin^2\phi + \beta_M \bar{\gamma}_s} \exp\left[-\frac{K\beta_M \bar{\gamma}_s}{(1+K)\sin^2\phi + \beta_M \bar{\gamma}_s}\right]. \tag{6.88}$$

■ Nakagami-m:

$$\mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) = \left(1 + \frac{\beta_M \bar{\gamma}_s}{m\sin^2\phi}\right)^{-m}. \tag{6.89}$$

All of these functions are simple trigonometrics and are therefore easy to integrate over the finite range in (6.85) or (6.86).

Example 6.5: Use the MGF technique to find an expression for the average probability of error for BPSK modulation in Nakagami fading.
Solution: We use the fact that BPSK for an AWGN channel has $P_b = Q(\sqrt{2\gamma_b})$, so $\alpha = 1$ and $\beta = 2$ in (6.83). The moment generating function for Nakagami-*m* fading is given by (6.89), and substituting this into (6.85) with $\alpha = 1$ and $\beta = 2$ yields

$$\bar{P}_b = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{\bar{\gamma}_b}{m\sin^2\phi}\right)^{-m} d\phi. \tag{6.90}$$

From (6.29) we see that the exact probability of symbol error for MQAM in AWGN contains both the Q-function and its square. Fortunately, an alternate form of $Q^2(z)$ allows us to apply the same techniques used here for MPSK to MQAM modulation. Specifically, an alternate representation of $Q^2(z)$ is derived in [15] as

$$Q^2(z) = \frac{1}{\pi} \int_0^{\pi/4} \exp\left[\frac{-z^2}{2\sin^2\phi}\right] d\phi. \tag{6.91}$$

Note that this is identical to the alternate representation for $Q(z)$ given in (6.53) except that the upper limit of the integral is $\pi/4$ instead of $\pi/2$. Thus we can write (6.29) in terms of the alternate representations for $Q(z)$ and $Q^2(z)$ as

$$P_s(\gamma_s) = \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)\int_0^{\pi/2} \exp\left(-\frac{\beta_M \gamma_s}{\sin^2\phi}\right) d\phi - \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)^2 \int_0^{\pi/4} \exp\left[-\frac{\beta_M \gamma_s}{\sin^2\phi}\right] d\phi, \tag{6.92}$$

where $\beta_M = 1.5/(M-1)$ is a function of the MQAM constellation size. Then the average probability of symbol error in fading becomes

$$\begin{aligned}
\bar{P}_s &= \int_0^\infty P_s(\gamma) p_{\gamma_s}(\gamma) d\gamma \\
&= \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)\int_0^{\pi/2}\int_0^\infty \exp\left[-\frac{\beta_M \gamma}{\sin^2\phi}\right] p_{\gamma_s}(\gamma)\, d\gamma\, d\phi \\
&\quad - \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)^2 \int_0^{\pi/4}\int_0^\infty \exp\left[-\frac{\beta_M \gamma}{\sin^2\phi}\right] p_{\gamma_s}(\gamma)\, d\gamma\, d\phi \\
&= \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)\int_0^{\pi/2} \mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) d\phi \\
&\quad - \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{M}}\right)^2 \int_0^{\pi/4} \mathcal{M}_{\gamma_s}\left(-\frac{\beta_M}{\sin^2\phi}\right) d\phi.
\end{aligned} \tag{6.93}$$

Thus, the exact average probability of symbol error is obtained via two finite-range integrals of the MGF over the fading distribution, which typically can be found in closed form or easily evaluated numerically.

The MGF approach can also be applied to noncoherent and differential modulations. For example, consider noncoherent MFSK, with P_s in AWGN given by (6.40), which is a finite sum of the desired form (6.78). Thus, in fading, the average symbol error probability of noncoherent MFSK is given by

$$\begin{aligned}
\bar{P}_s &= \int_0^\infty \sum_{m=1}^{M} (-1)^{m+1}\binom{M-1}{m}\frac{1}{m+1}\exp\left[\frac{-m\gamma}{m+1}\right] p_{\gamma_s}(\gamma)\, d\gamma \\
&= \sum_{m=1}^{M} (-1)^{m+1}\binom{M-1}{m}\frac{1}{m+1}\left[\int_0^\infty \exp\left[\frac{-m\gamma}{m+1}\right] p_{\gamma_s}(\gamma)\, d\gamma\right] \\
&= \sum_{m=1}^{M} (-1)^{m+1}\binom{M-1}{m}\frac{1}{m+1}\mathcal{M}_{\gamma_s}\left(-\frac{m}{m+1}\right).
\end{aligned} \tag{6.94}$$

Finally, for differential MPSK it can be shown [16] that the average probability of symbol error is given by

$$P_s = \frac{\sqrt{\beta_M}}{2\pi}\int_{-\pi/2}^{\pi/2} \frac{\exp[-\gamma_s(1-\sqrt{1-\beta_M}\cos\theta)]}{1-\sqrt{1-\beta_M}\cos\theta}\, d\theta \tag{6.95}$$

for $\beta_M = \sin^2(\pi/M)$, which is in the desired form (6.79). Thus we can express the average probability of symbol error in terms of the MGF of the fading distribution as

$$\bar{P}_s = \frac{\sqrt{\beta_M}}{2\pi} \int_{-\pi/2}^{\pi/2} \frac{\mathcal{M}_{\gamma s}(-(1-\sqrt{1-\beta_M}\cos\theta))}{1-\sqrt{1-\beta_M}\cos\theta}\, d\theta. \tag{6.96}$$

A more extensive discussion of the MGF technique for finding average probability of symbol error for different modulations and fading distributions can be found in [9, Chap. 8.2].

6.3.4 Combined Outage and Average Error Probability

When the fading environment is a superposition of both fast and slow fading (e.g., log-normal shadowing and Rayleigh fading), a common performance metric is combined outage and average error probability, where outage occurs when the slow fading falls below some target value and the average performance in non-outage is obtained by averaging over the fast fading. We use the following notation.

- $\bar{\bar{\gamma}}_s$ denotes the average SNR per symbol for a fixed path loss with averaging over fast fading and shadowing.
- $\bar{\gamma}_s$ denotes the (random) SNR per symbol for a fixed path loss and random shadowing but averaged over fast fading. Its average value, averaged over the shadowing, is $\bar{\bar{\gamma}}_s$.
- γ_s denotes the random SNR due to fixed path loss, shadowing, and multipath. Its average value, averaged over multipath only, is $\bar{\gamma}_s$. Its average value, averaged over both multipath and shadowing, is $\bar{\bar{\gamma}}_s$.

With this notation we can specify an average error probability $\bar{P}_s$ with some probability $1 - P_{\text{out}}$. An outage is declared when the received SNR per symbol due to shadowing and path loss alone, $\bar{\gamma}_s$, falls below a given target value $\bar{\gamma}_{s_0}$. When not in outage ($\bar{\gamma}_s \geq \bar{\gamma}_{s_0}$), the average probability of error is obtained by averaging over the distribution of the fast fading conditioned on the mean SNR:

$$\bar{P}_s = \int_0^\infty P_s(\gamma_s) p(\gamma_s \mid \bar{\gamma}_s)\, d\gamma_s. \tag{6.97}$$

The criterion used to determine the outage target $\bar{\gamma}_{s_0}$ is typically based on a given maximum acceptable average probability of error $\bar{P}_{s_0}$. The target $\bar{\gamma}_{s_0}$ must then satisfy

$$\bar{P}_{s_0} = \int_0^\infty P_s(\gamma_s) p(\gamma_s \mid \bar{\gamma}_{s_0})\, d\gamma_s. \tag{6.98}$$

It is clear that, whenever $\bar{\gamma}_s > \bar{\gamma}_{s_0}$, the average error probability $\bar{P}_s$ will be below the target maximum value $\bar{P}_{s_0}$.

Example 6.6: Consider BPSK modulation in a channel with both log-normal shadowing ($\sigma_{\psi_{\text{dB}}} = 8\,\text{dB}$) and Rayleigh fading. The desired maximum average error probability is $\bar{P}_{b_0} = 10^{-4}$, which requires $\bar{\gamma}_{b_0} = 34\,\text{dB}$. Determine the value of $\bar{\bar{\gamma}}_b$ that will ensure $\bar{P}_b \leq 10^{-4}$ with probability $1 - P_{\text{out}} = .95$.

Solution: We must find $\bar{\bar{\gamma}}_b$, the average of γ_b in both the fast and slow fading, such that $p(\bar{\gamma}_b > \bar{\gamma}_{b_0}) = 1 - P_{\text{out}}$. For log-normal shadowing we compute this as

$$p(\bar{\gamma}_b > 34) = p\left(\frac{\bar{\gamma}_b - \bar{\bar{\gamma}}_b}{\sigma_{\psi_{\text{dB}}}} \geq \frac{34 - \bar{\bar{\gamma}}_b}{\sigma_{\psi_{\text{dB}}}}\right) = Q\left(\frac{34 - \bar{\bar{\gamma}}_b}{\sigma_{\psi_{\text{dB}}}}\right) = 1 - P_{\text{out}}, \tag{6.99}$$

since, assuming units in dB, $(\bar{\gamma}_b - \bar{\bar{\gamma}}_b)/\sigma_{\psi_{\text{dB}}}$ is a Gauss-distributed random variable with mean 0 and standard deviation 1. Thus, the value of $\bar{\bar{\gamma}}_b$ is obtained by substituting the values of P_{out} and $\sigma_{\psi_{\text{dB}}}$ in (6.99) and using a table of Q-functions or an inversion program, which yields $(34 - \bar{\bar{\gamma}}_b)/8 = -1.6$ or $\bar{\bar{\gamma}}_b = 46.8\,\text{dB}$.

6.4 Doppler Spread

One consequence of Doppler spread is an irreducible error floor for modulation techniques using differential detection. This is due to the fact that in differential modulation the signal phase associated with one symbol is used as a phase reference for the next symbol. If the channel phase decorrelates over a symbol, then the phase reference becomes extremely noisy, leading to a high symbol error rate that is independent of received signal power. The phase correlation between symbols and consequent degradation in performance are functions of the Doppler frequency $f_D = v/\lambda$ and the symbol time T_s.

The first analysis of the irreducible error floor due to Doppler was done by Bello and Nelin in [17]. In that work, analytical expressions for the irreducible error floor of noncoherent FSK and DPSK due to Doppler are determined for a Gaussian Doppler power spectrum. However, these expressions are not in closed form, so they must be evaluated numerically. Closed-form expressions for the bit error probability of DPSK in fast Rician fading, where the channel decorrelates over a bit time, can be obtained using the MGF technique, with the MGF obtained based on the general quadratic form of complex Gaussian random variables (see [1, Apx. B] and [12, Apx. B]). A different approach utilizing alternate forms of the Marcum Q-function can also be used [9, Chap. 8.2.5]. The resulting average bit error probability for DPSK is

$$\bar{P}_b = \frac{1}{2}\left(\frac{1+K+\bar{\gamma}_b(1-\rho_C)}{1+K+\bar{\gamma}_b}\right)\exp\left[-\frac{K\bar{\gamma}_b}{1+K+\bar{\gamma}_b}\right], \tag{6.100}$$

where ρ_C is the channel correlation coefficient after a bit time T_b, K is the fading parameter of the Rician distribution, and $\bar{\gamma}_b$ is the average SNR per bit. For Rayleigh fading ($K = 0$) this simplifies to

$$\bar{P}_b = \frac{1}{2}\left(\frac{1+\bar{\gamma}_b(1-\rho_C)}{1+\bar{\gamma}_b}\right). \tag{6.101}$$

Letting $\bar{\gamma}_b \to \infty$ in (6.100) yields the irreducible error floor:

$$\bar{P}_{\text{floor}} = \frac{(1-\rho_C)e^{-K}}{2} \text{ (DPSK)}. \tag{6.102}$$

A similar approach is used in [18] to bound the bit error probability of DQPSK in fast Rician fading as

$$\bar{P}_b \le \frac{1}{2}\left(1-\sqrt{\frac{(\rho_C\bar{\gamma}_s/\sqrt{2})^2}{(\bar{\gamma}_s+1)^2-(\rho_C\bar{\gamma}_s/\sqrt{2})^2}}\right)\exp\left[-\frac{(2-\sqrt{2})K\bar{\gamma}_s/2}{(\bar{\gamma}_s+1)-(\rho_C\bar{\gamma}_s/\sqrt{2})}\right], \tag{6.103}$$

where K is as before, ρ_C is the channel correlation coefficient after a symbol time T_s, and $\bar{\gamma}_s$ is the average SNR per symbol. Letting $\bar{\gamma}_s \to \infty$ yields the irreducible error floor:

$$\bar{P}_{\text{floor}} = \frac{1}{2}\left(1-\sqrt{\frac{(\rho_C/\sqrt{2})^2}{1-(\rho_C/\sqrt{2})^2}}\right)\exp\left[-\frac{(2-\sqrt{2})(K/2)}{1-\rho_C/\sqrt{2}}\right] \text{ (DQPSK)}. \tag{6.104}$$

As discussed in Section 3.2.1, the channel correlation $A_C(\tau)$ over time τ equals the inverse Fourier transform of the Doppler power spectrum $S_C(f)$ as a function of Doppler frequency f. The correlation coefficient is thus $\rho_C = A_C(T)/A_C(0)$ evaluated at $T = T_s$ for DQPSK or at $T = T_b$ for DPSK. Table 6.2, from [19], gives the value of ρ_C for several different Doppler power spectral models, where B_D is the Doppler spread of the channel. Assuming the uniform scattering model ($\rho_C = J_0(2\pi f_D T_b)$) and Rayleigh fading ($K = 0$) in (6.102) yields an irreducible error for DPSK of

$$\overline{P}_{\text{floor}} = \frac{1-J_0(2\pi f_D T_b)}{2} \approx .5(\pi f_D T_b)^2, \tag{6.105}$$

Table 6.2: Correlation coefficients for different Doppler power spectra models

Type	**Doppler power spectrum $S_C(f)$**	$\rho_C = A_C(T)/A_C(0)$
Rectangular	$P_0/(2B_D), \|f\| < B_D$	$\mathrm{sinc}(2B_DT)$
Gaussian	$(P_0/(\sqrt{\pi}B_D))e^{-f^2/B_D^2}$	$e^{-(\pi B_D T)^2}$
Uniform scattering	$P_0/\left(\pi\sqrt{B_D^2 - f^2}\right),\ \|f\| < B_D$	$J_0(2\pi B_D T)$
1st-order Butterworth	$P_0 B_D/\left(\pi(f^2 + B_D^2)\right)$	$e^{-2\pi B_D T}$

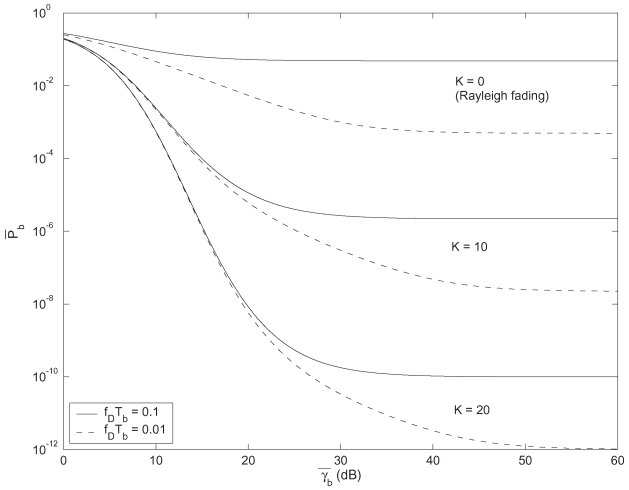


Figure 6.5: Average P_b for DPSK in fast Rician fading with uniform scattering.

where $B_D = f_D = v/\lambda$ is the maximum Doppler in the channel. Note that in this expression the error floor decreases with data rate $R = 1/T_b$. This is true in general for irreducible error floors of differential modulation due to Doppler, since the channel has less time to decorrelate between transmitted symbols. This phenomenon is one of the few instances in digital communications where performance improves as data rate increases.

A plot of (6.100), the error probability of DPSK in fast Rician fading, for uniform scattering ($\rho_C = J_0(2\pi f_D T_b)$) and different values of $f_D T_b$ is shown in Figure 6.5. We see from this figure that the error floor starts to dominate at $\overline{\gamma}_b = 15$ dB in Rayleigh fading ($K = 0$), and as K increases the value of $\overline{\gamma}_b$ where the error floor dominates also increases. We also see that increasing the data rate $R_b = 1/T_b$ by an order of magnitude decreases the error floor by roughly two orders of magnitude.

Example 6.7: Assume a Rayleigh fading channel with uniform scattering and a maximum Doppler of $f_D = 80$ Hz. For what approximate range of data rates will the irreducible error floor of DPSK be below 10^{-4}?
Solution: We have $\overline{P}_{\text{floor}} \approx .5(\pi f_D T_b)^2 < 10^{-4}$. Solving for T_b with $f_D = 80$ Hz, we get

$$T_b < \frac{\sqrt{2 \cdot 10^{-4}}}{\pi \cdot 80} = 5.63 \cdot 10^{-5}, \tag{6.106}$$

which yields $R > 17.77$ kbps.

Deriving analytical expressions for the irreducible error floor becomes intractable with more complex modulations, in which case simulations are often used. In particular, simulations of the irreducible error floor for $\pi/4$-DQPSK with square-root raised cosine filtering have been conducted (since this modulation is used in the IS-136 TDMA standard) in [20, 21]. These simulation results indicate error floors between 10^{-3} and 10^{-4}. As expected, in these simulations the error floor increases with vehicle speed, since at higher vehicle speeds the channel typically decorrelates more over a given symbol time.

6.5 Intersymbol Interference

Frequency-selective fading gives rise to intersymbol interference, where the received symbol over a given symbol period experiences interference from other symbols that have been delayed by multipath. Since increasing signal power also increases the power of the ISI, this interference gives rise to an irreducible error floor that is independent of signal power. The irreducible error floor is difficult to analyze because it depends on the modulation format and the ISI characteristics, which in turn depend on the characteristics of the channel and the sequence of transmitted symbols.

The first extensive analysis of the degradation in symbol error probability due to ISI was done by Bello and Nelin [22]. In that work, analytical expressions for the irreducible error floor of coherent FSK and noncoherent DPSK are determined assuming a Gaussian delay profile for the channel. To simplify the analysis, only ISI associated with adjacent symbols was taken into account. Even with this simplification, the expressions are complex and must be approximated for evaluation. A more accurate analytical analysis can be found in [23, Chap. 8.2], where the irreducible error floor is evaluated based on both the worst-case sequence of transmitted symbols and by averaging over all possible symbol sequences. These expressions are also complex to evaluate owing to their dependence on the channel and symbol sequence characteristics. A simple approximation to symbol error probability with ISI can be obtained by treating the ISI as uncorrelated white Gaussian noise. Then the SNR becomes

$$\hat{\gamma}_s = \frac{P_r}{N_0 B + I}, \tag{6.107}$$

where P_r is the received power associated with the line-of-sight signal component, and I is the received power associated with the ISI. In a static channel the resulting probability of symbol error will be $P_s(\hat{\gamma}_s)$, where P_s is the probability of symbol error in AWGN. If both the LOS and ISI signal components experience flat fading, then $\hat{\gamma}_s$ will be a random variable with distribution $p(\hat{\gamma}_s)$. The outage probability is then $P_{\text{out}} = p(\hat{\gamma}_s < \gamma_0)$ and the average symbol error probability is $\overline{P}_s = \int P_s(\hat{\gamma}_s) p(\hat{\gamma}_s)\, d\hat{\gamma}_s$. Note that $\hat{\gamma}_s$ is the ratio of two random variables – the LOS received power P_r and the ISI received power I – and thus the resulting distribution $p(\hat{\gamma}_s)$ may be hard to obtain and typically is not in closed form.

Irreducible error floors due to ISI are often obtained by simulation, which can easily incorporate different channel models, modulation formats, and symbol sequence characteristics [24, 20, 21, 25, 26, 27]. The earliest of these works is [24]. The extensive simulations in this work determined the irreducible error floor due to ISI for BPSK, DPSK, QPSK, OQPSK and MSK modulations with different pulse shapes and for channels with different power delay profiles, including a Gaussian, exponential, equal-amplitude two-ray, and empirical power delay profile. The results of [24] indicate that the irreducible error floor is more sensitive to the rms delay spread of the channel than to the shape of its power delay profile. Moreover, pulse shaping can significantly impact the error floor: for the raised cosine pulses discussed in Section 5.5, increasing β from 0 to 1 can reduce the error floor by over an order of magnitude.

The simulation results of [24] are shown in Figure 6.6. This figure plots the irreducible bit error rate as a function of normalized rms delay spread $d = \sigma_{T_m}/T_s$ for BPSK, QPSK, OQPSK, and MSK modulation assuming a static channel with a Gaussian power delay profile. We see from the figure that, for all modulations, we can

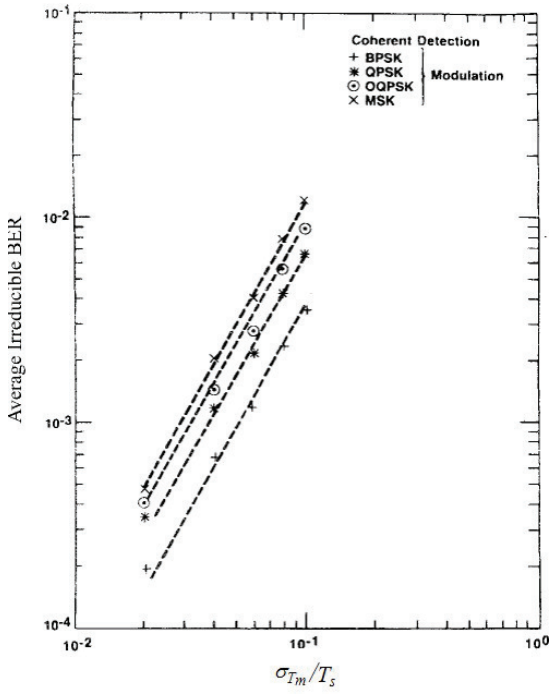


Figure 6.6: Irreducible error versus normalized rms delay spread for Gaussian power delay profile. (Reprinted by permission from [24, Fig. 9], © 1987 IEEE.)

approximately bound the irreducible error floor as

$$P_{\text{floor}} \leq K \left(\frac{\sigma_{T_m}}{T_s} \right)^2 \tag{6.108}$$

for K a constant that depends on the modulation type, the transmit and receive filtering, and the power delay profile. For the modulation types in Figure 6.6 we see that K ranges from about 0.4 for BPSK to 1 for MSK. Other simulation results [27, 25, 26] support this bound as well. This bound can also be derived analytically based on group delay. Details can be found in [29, Chapter 12.3.2].

This bound imposes severe constraints on data rate even when symbol error probabilities on the order of 10^{-2} are acceptable, and usually this is a very high error probability even with strong error correction codes. For example, the rms delay spread in a typical urban environment is approximately $\sigma_{T_m} = 2.5$ μs. To keep $\sigma_{T_m} < .1T_s$ requires that the data rate not exceed 40 kbaud, which generally isn't enough for high-speed data applications. In rural environments, where multipath is not attenuated to the same degree as in cities, $\sigma_{T_m} \approx 25$ μs, which reduces the maximum data rate to 4 kbaud.

Example 6.8: Using the approximation $\overline{P}_{\text{floor}} \leq K(\sigma_{T_m}/T_s)^2$ of (6.108) with $K = 0.4$ for BPSK and $K = 0.7$ for QPSK, find the maximum data rate that can be transmitted through a channel with delay spread $\sigma_{T_m} = 3$ μs, using either BPSK or QPSK modulation, such that the probability of bit error P_b is less than 10^{-3}.

Solution: For BPSK, we set $\overline{P}_{\text{floor}} = .4(\sigma_{T_m}/T_b)^2$ and so require $T_b \geq \sigma_{T_m}/\sqrt{\overline{P}_{\text{floor}}/.4} = 60$ μs, which leads to a data rate of $R = 1/T_b = 16.67$ Kbps. For QPSK, we have $T_s \geq \sigma_{T_m}/\sqrt{\overline{P}_{\text{floor}}/.7} = 79$ μs. Since there are two bits per symbol, this leads to a data rate of $R = 2/T_s = 25.2$ Kbps. This indicates that, for a given data rate,

QPSK is more robust to ISI than BPSK because its symbol time is slower, even though the K value in the error floor bound (6.108) is higher for the larger constellation size of QPSK versus BPSK. In general K increases in the bound of (6.108) as the constellation size increases, but large constellations are still more robust to ISI than smaller ones due to their longer symbol times.

Chapter 6 Problems

1. Consider a system in which data is transferred at a rate of 100 bits per second over the channel.

   (a) Find the symbol duration if we use a sinc pulse for signaling and the channel bandwidth is 10 kHz.

   (b) Suppose the received SNR is 10 dB. Find the SNR per symbol and the SNR per bit if 4-QAM is used.

   (c) Find the SNR per symbol and the SNR per bit for 16-QAM, and compare with these metrics for 4-QAM.

2. Consider BPSK modulation where the a priori probability of 0 and 1 is not the same. Specifically, $p(s_n = 0) = 0.3$ and $p(s_n = 1) = 0.7$.

   (a) Find the probability of bit error P_b in AWGN assuming we encode a 1 as $s_1(t) = A\cos(2\pi f_c t)$ and a 0 as $s_2(t) = -A\cos(2\pi f_c t)$ for $A > 0$, assuming the receiver structure is as shown in Figure 5.17.

   (b) Suppose you can change the threshold value in the receiver of Figure 5.17. Find the threshold value that yields equal error probability regardless of which bit is transmitted – that is, the threshold value that yields $p(\hat{m} = 0 \,|\, m = 1)p(m = 1) = p(\hat{m} = 1 \,|\, m = 0)p(m = 0)$.

   (c) Now suppose we change the modulation so that $s_1(t) = A\cos(2\pi f_c t)$ and $s_2(t) = -B\cos(2\pi f_c t)$. Find $A > 0$ and $B > 0$ so that the receiver of Figure 5.17 with threshold at zero has $p(\hat{m} = 0 \,|\, m = 1)p(m = 1) = p(\hat{m} = 1 \,|\, m = 0)p(m = 0)$.

   (d) Compute and compare the expression for P_b in parts (a), (b), and (c) assuming $E_b/N_0 = 10$ dB and $N_0 = .1$. For which system is P_b minimized?

3. Consider a BPSK receiver whose demodulator has a phase offset of ϕ relative to the transmitted signal, so for a transmitted signal $s(t) = \pm g(t)\cos(2\pi f_c t)$ the carrier in the demodulator of Figure 5.17 is $\cos(2\pi f_c t + \phi)$. Determine the threshold level in the threshold device of Figure 5.17 that minimizes probability of bit error, and find this minimum error probability.

4. Assume a BPSK demodulator in which the receiver noise is added after the integrator, as shown in Figure 6.7. The decision device outputs a 1 if its input x has $\mathrm{Re}\{x\} \geq 0$, and a 0 otherwise. Suppose the tone jammer $n(t) = 1.1e^{j\theta}$, where $p(\theta = n\pi/3) = 1/6$ for $n = 0, 1, 2, 3, 4, 5$. What is the probability of making a decision error in the decision device (i.e., outputting the wrong demodulated bit), assuming $A_c = \sqrt{2/T_b} = 1$ and that information bits corresponding to a 1 ($s(t) = A_c\cos(2\pi f_c t)$) or a 0($s(t) = -A_c\cos(2\pi f_c t)$) are equally likely.

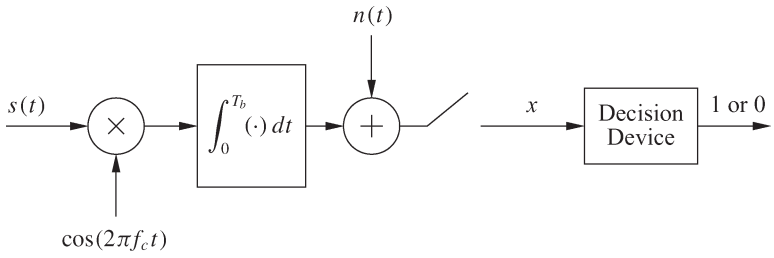


Figure 6.7: BPSK demodulator for Problem 6-4.

5. Find an approximation to P_s for the signal constellations shown in Figure 6.8.

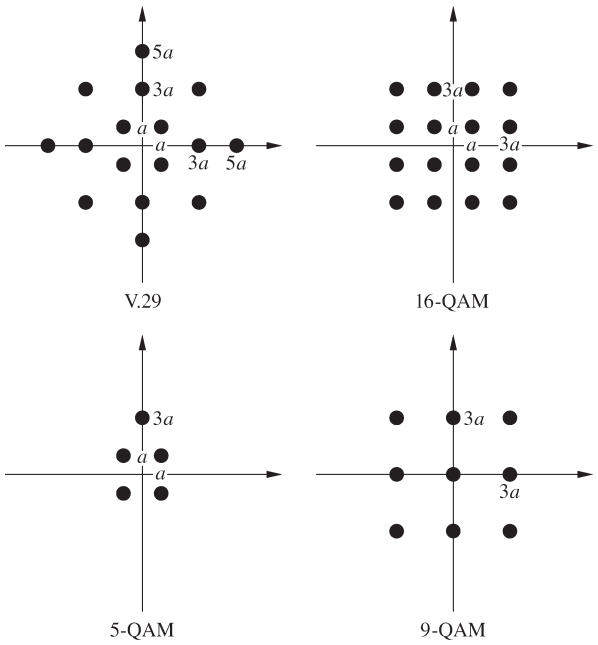


Figure 6.8: Signal constellations for Problem 6-5.

6. Plot the exact symbol error probability and the approximation from Table 6.1 of 16-QAM with $0 \leq \gamma_s \leq 30$ dB. Does the error in the approximation increase or decrease with γ_s? Why?

7. Plot the symbol error probability P_s for QPSK using the approximation in Table 6.1 and Craig's exact result for $0 \leq \gamma_s \leq 30$ dB. Does the error in the approximation increase or decrease with γ_s? Why?

8. In this problem we derive an algebraic proof of the alternate representation of the Q-function (6.53) from its original representation (6.52). We will work with the complementary error function (erfc) for simplicity and make the conversion at the end. The erfc(x) function is traditionally defined by

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \tag{6.109}$$

The alternate representation of this, corresponding to the alternate representation of the Q-function (6.53), is

$$\text{erfc}(x) = \frac{2}{\pi} \int_0^{\pi/2} e^{-x^2/\sin^2\theta} d\theta. \tag{6.110}$$

(a) Consider the integral

$$I_x(a) \triangleq \int_0^\infty \frac{e^{-at^2}}{x^2 + t^2} dt, \tag{6.111}$$

and show that $I_x(a)$ satisfies the following differential equation:

$$x^2 I_x(a) - \frac{\partial I_x(a)}{\partial a} = \frac{1}{2}\sqrt{\frac{\pi}{a}}. \tag{6.112}$$

(b) Solve the differential equation (6.112) and deduce that

$$I_x(a) \triangleq \int_0^\infty \frac{e^{-at^2}}{x^2+t^2}\, dt = \frac{\pi}{2x} e^{ax^2} \operatorname{erfc}(x\sqrt{a}). \tag{6.113}$$

Hint: $I_x(a)$ is a function in two variables x and a. However, since all our manipulations deal with a only, you can assume x to be a constant while solving the differential equation.

(c) Setting $a = 1$ in (6.113) and making a suitable change of variables in the left-hand side of (6.113), derive the alternate representation of the erfc function:

$$\operatorname{erfc}(x) = \frac{2}{\pi} \int_0^{\pi/2} e^{-x^2/\sin^2\theta} d\theta.$$

(d) Convert this alternate representation of the erfc function to the alternate representation of the Q-function.

9. Consider a communication system that uses BPSK signaling, with average signal power of 100 W and noise power at the receiver of 4 W. Based on its error probability, can this system be used for transmission of data? Can it be used for voice? Now consider the presence of fading with an average SNR $\overline{\gamma}_b = 20$ dB. How do your answers to the previous questions change?

10. Consider a cellular system at 900 MHz with a transmission rate of 64 kbps and multipath fading. Explain which performance metric – average probability of error or outage probability – is more appropriate (and why) for user speeds of l mph, 10 mph, and 100 mph.

11. Derive the expression for the moment generating function for SNR in Rayleigh fading.

12. This problem illustrates why satellite systems that must compensate for shadow fading are going bankrupt. Consider an LEO satellite system orbiting 500 km above the earth. Assume that the signal follows a free-space path-loss model with no multipath fading or shadowing. The transmitted signal has a carrier frequency of 900 MHz and a bandwidth of 10 kHz. The handheld receivers have noise power spectral density of 10^{-16} mW/Hz (total noise power is N_0B). Assume nondirectional antennas (0-dB gain) at both the transmitter and receiver. Suppose the satellite must support users in a circular cell on the earth of radius 100 km at a BER of 10^{-6}.

(a) For DPSK modulation, find the transmit power needed for all users in the cell to meet the 10^{-6} BER target.

(b) Repeat part (a) assuming that the channel also experiences log-normal shadowing with $\sigma_{\psi_{\text{dB}}} = 8$ dB and that users in a cell must have $P_b = 10^{-6}$ (for each bit) with probability 0.9.

13. In this problem we explore the power penalty involved in going from BPSK to the higher-level signal modulation of 16-PSK.

(a) Find the minimum distance between constellation points in 16-PSK modulation as a function of signal energy E_s.

(b) Find α_M and β_M such that the symbol error probability of 16-PSK in AWGN is approximately

$$P_s \approx \alpha_M Q(\sqrt{\beta_M \gamma_s}).$$

(c) Using your expression in part (b), find an approximation for the average symbol error probability of 16-PSK in Rayleigh fading in terms of $\overline{\gamma}_s$.

(d) Convert the expressions for average symbol error probability of 16-PSK in Rayleigh fading to an expression for average bit error probability, assuming Gray coding.

(e) Find the approximate value of $\overline{\gamma}_b$ required to obtain a BER of 10^{-3} in Rayleigh fading for BPSK and 16-PSK. What is the power penalty in going to the higher-level signal constellation at this BER?

14. Find a closed-form expression for the average probability of error for DPSK modulation in Nakagami-m fading. Evaluate for $m = 4$ and $\overline{\gamma}_b = 10$ dB.

15. The Nakagami distribution is parameterized by m, which ranges from $m = .5$ to $m = \infty$. The m-parameter measures the ratio of LOS signal power to multipath power, so $m = 1$ corresponds to Rayleigh fading, $m = \infty$ corresponds to an AWGN channel with no fading, and $m = .5$ corresponds to fading that results in performance that is worse than with a Rayleigh distribution. In this problem we explore the impact of the parameter m on the performance of BPSK modulation in Nakagami fading.

    Plot the average bit error $\overline{P}_b$ of BPSK modulation in Nakagami fading with average SNR ranging from 0 dB to 20 dB for m parameters $m = 1$ (Rayleigh), $m = 2$, and $m = 4$. (The moment generating function technique of Section 6.3.3 should be used to obtain the average error probability.) At an average SNR of 10 dB, what is the difference in average BER?

16. Assume a cellular system with log-normal shadowing plus Rayleigh fading. The signal modulation is DPSK. The service provider has determined that it can deal with an outage probability of .01 – that is, 1 in 100 customers can be unhappy at any given time. In nonoutage, the voice BER requirement is $\overline{P}_b = 10^{-3}$. Assume a noise PSD $N_0/2$ with $N_0 = 10^{-16}$ mW/Hz, a signal bandwidth of 30 kHz, a carrier frequency of 900 MHz, free-space path-loss propagation with nondirectional antennas, and a shadowing standard deviation of $\sigma_{\psi_{\rm dB}} = 6$ dB. Find the maximum cell size that can achieve this performance if the transmit power at the mobiles is limited to 100 mW.

17. Consider a cellular system with circular cells of radius 100 meters. Assume that propagation follows the simplified path-loss model with $K = 1, d_0 = 1$ m, and $\gamma = 3$. Assume the signal experiences (in addition to path loss) log-normal shadowing with $\sigma_{\psi_{\rm dB}} = 4$ as well as Rayleigh fading. The transmit power at the base station is $P_t = 100$ mW, the system bandwidth is $B = 30$ kHz, and the noise PSD $N_0/2$ has $N_0 = 10^{-14}$W/Hz. Assuming BPSK modulation, we want to find the cell coverage area (percentage of locations in the cell) where users have average P_b of less than 10^{-4}.

    (a) Find the received power due to path loss at the cell boundary.

    (b) Find the minimum average received power (due to path loss and shadowing) such that, with Rayleigh fading about this average, a BPSK modulated signal with this average received power at a given cell location has $\overline{P}_b < 10^{-4}$.

    (c) Given the propagation model for this system (simplified path loss, shadowing, and Rayleigh fading), find the percentage of locations in the cell where $\overline{P}_b < 10^{-4}$ under BPSK modulation.

18. In this problem we derive the probability of bit error for DPSK in fast Rayleigh fading. By symmetry, the probability of error is the same for transmitting a 0-bit or a 1-bit. Let us assume that over time kT_b a 0-bit is transmitted, so the transmitted symbol at time kT_b is the same as at time $k-1$: $\mathbf{s}(k) = \mathbf{s}(k-1)$. In fast fading, the corresponding received symbols are $\mathbf{r}(k-1) = g_{k-1}\mathbf{s}(k-1) + n(k-1)$ and $\mathbf{r}(k) = g_k\mathbf{s}(k-1) + n(k)$, where g_{k-1} and g_k are the fading channel gains associated with transmissions over times $(k-1)T_b$ and kT_b.

(a) Show that the decision variable that is input to the phase comparator of Figure 5.20 in order to extract the phase difference is $\mathbf{r}(k)\mathbf{r}^*(k-1) = g_k g_{k-1}^* + g_k \mathbf{s}(k-1)^* n_{k-1} + g_{k-1}^* s_{k-1}^* n_k + n_k n_{k-1}^*$.

Assuming a reasonable SNR, the last term $n_k n_{k-1}^*$ of this expression can be neglected. So neglecting this term and then defining $\tilde{n}_k = s_{k-1}^* n_k$ and $\tilde{n}_{k-1} = s_{k-1}^* n_{k-1}$, we get a new random variable $\tilde{z} = g_k g_{k-1}^* + g_k \tilde{n}_{k-1}^* + g_{k-1}^* \tilde{n}_k$. Given that a 0-bit was transmitted over time kT_b, an error is made if $x = \mathrm{Re}\{\tilde{z}\} < 0$, so we must determine the distribution of x. The characteristic function for x is the two-sided Laplace transform of the distribution of x:

$$\Phi_X(s) = \int_{-\infty}^{\infty} p_X(s) e^{-sx}\, dx = \mathbf{E}[e^{-sx}].$$

This function will have a left-plane pole p_1 and a right-plane pole p_2, so it can be written as

$$\Phi_X(s) = \frac{p_1 p_2}{(s-p_1)(s-p_2)}.$$

The left-plane pole p_1 corresponds to the distribution $p_X(x)$ for $x \geq 0$, and the rightplane pole corresponds to the distribution $p_X(x)$ for $x < 0$.

(b) Show through partial fraction expansion that $\Phi_X(s)$ can be written as

$$\Phi_X(s) = \frac{p_1 p_2}{p_1 - p_2} \frac{1}{s - p_1} + \frac{p_1 p_2}{p_2 - p_1} \frac{1}{s - p_2}.$$

An error is made if $x = \mathrm{Re}\{\tilde{z}\} < 0$, so here we need only consider the distribution $p_X(x)$ for $x < 0$ corresponding to the second term of $\Phi_X(s)$.

(c) Show that the inverse Laplace transform of the second term of $\Phi_X(s)$ from part (b) is

$$p_X(x) = \frac{p_1 p_2}{p_2 - p_1} e^{p_2 X},\ x < 0.$$

(d) Use part (c) to show that $P_b = -p_1/(p_2 - p_1)$.

In $x = \mathrm{Re}\{\tilde{z}\} = \mathrm{Re}\{g_k g_{k-1}^* + g_k \tilde{n}_{k-1}^* + g_{k-1}^* \tilde{n}_k\}$, the channel gains g_k, g_{k-1} and noises $\tilde{n}_k, \tilde{n}_{k-1}$ are complex Gaussian random variables. Thus, the poles p_1, p_2 in $p_X(x)$ are derived using the general quadratic form of complex Gaussian random variables (see [1, Apx. B] and [12, Apx. B]) as

$$p_1 = \frac{-1}{N_0(\overline{\gamma}_b[1+\rho_c]+1)} \text{ and } p_2 = \frac{1}{N_0(\overline{\gamma}_b[1-\rho_c]+1)}$$

for ρ_C the correlation coefficient of the channel over the bit time T_b.

(e) Find a general expression for P_b in fast Rayleigh fading using these values of p_1 and p_2 in the P_b expression from part (d).

(f) Show that this reduces to the average probability of error $\overline{P}_b = 1/2(1+\overline{\gamma}_b)$ for a slowly fading channel that does not decorrelate over a bit time.

19. Plot the bit error probability for DPSK in fast Rayleigh fading for $\overline{\gamma}_b$ ranging from 0 dB to 60 dB and $\rho_C = J_0(2\pi B_D T)$ with $B_D T = .01, .001$, and $.0001$. For each value of $B_D T$, at approximately what value of $\overline{\gamma}_b$ does the error floor dominate the error probability?

20. Find the irreducible error floor due to Doppler for DQPSK modulation with a data rate of 40 kbps, assuming a Gaussian Doppler power spectrum with $B_D = 80$ Hz and Rician fading with $K = 2$.

21. Consider a wireless channel with an average delay spread of 100 ns and a Doppler spread of 80 Hz. Given the error floors due to Doppler and ISI – and assuming DQPSK modulation in Rayleigh fading and uniform scattering – approximately what range of data rates can be transmitted over this channel with a BER of less than 10^{-4}?

22. Using the error floors of Figure 6.6, find the maximum data rate that can be transmitted through a channel with delay spread $\sigma_{T_m} = 3$ μs (using BPSK, QPSK, or MSK modulation) such that the probability of bit error P_b is less than 10^{-3}.

Bibliography

[1] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed., McGraw-Hill, New York, 2008.

[2] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[3] S. Rhodes, "Effect of noisy phase reference on coherent detection of offset-QPSK signals," *IEEE Trans. Commun.*, pp. 1046–55, August 1974.

[4] N. R. Sollenberger and J. C.-I. Chuang, "Low-overhead symbol timing and carrier recovery for portable TDMA radio systems," *IEEE Trans. Commun.*, pp. 1886–92, October 1990.

[5] R. Pawula, S. Rice, and J. Roberts, "Distribution of the phase angle between two vectors perturbed by Gaussian noise," *IEEE Trans. Commun.*, pp. 1828–41, August 1982.

[6] W. Cowley and L. Sabel, "The performance of two symbol timing recovery algorithms for PSK demodulators," *IEEE Trans. Commun.*, pp. 2345–55, June 1994.

[7] W. T. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation*, IEEE/Pentech Press, London, 1994.

[8] X. Tang, M.-S. Alouini, and A. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Trans. Commun.*, pp. 1856–64, December 1999.

[9] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, 2nd Ed., Wiley, New York, 2004.

[10] S. Hinedi, M. Simon, and D. Raphaeli, "The performance of noncoherent orthogonal M-FSK in the presence of timing and frequency errors," *IEEE Trans. Commun.*, pp. 922–33, February–April 1995.

[11] E. Grayver and B. Daneshrad, "A low-power all-digital FSK receiver for deep space applications," *IEEE Trans. Commun.*, pp. 911–21, May 2001.

[12] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*, McGraw-Hill, New York, 1966 [reprinted 1995 by Wiley/IEEE Press].

[13] J. Craig, "New, simple and exact result for calculating the probability of error for two-dimensional signal constellations," *Proc. Military Commun. Conf.*, pp. 25.5.1–25.5.5, November 1991.

[14] F. S. Weinstein, "Simplified relationships for the probability distribution of the phase of a sine wave in narrowband normal noise," *IEEE Trans. Inform. Theory*, pp. 658–61, September 1974.

[15] M. K. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Trans. Commun.*, pp. 200–10, February 1998.

[16] R. F. Pawula, "A new formula for MDPSK symbol error probability," *IEEE Commun. Lett*., pp. 271–2, October 1998.

[17] P. A. Bello and B. D. Nelin, "The influence of fading spectrum on the bit error probabilities of incoherent and differentially coherent matched filter receivers," *IEEE Trans. Commun. Syst*., pp. 160–8, June 1962.

[18] P. Y. Kam, "Tight bounds on the bit-error probabilities of 2DPSK and 4DPSK in nonselective Rician fading," *IEEE Trans. Commun*., pp. 860–2, July 1998.

[19] P. Y. Kam, "Bit error probabilities of MDPSK over the nonselective Rayleigh fading channel with diversity reception," *IEEE Trans. Commun*., pp. 220–4, February 1991.

[20] V. Fung, R. S. Rappaport, and B. Thoma, "Bit error simulation for $\pi/4$ DQPSK mobile radio communication using two-ray and measurement based impulse response models," *IEEE J. Sel. Areas Commun*., pp. 393–405, April 1993.

[21] S. Chennakeshu and G. J. Saulnier, "Differential detection of $\pi/4$-shifted-DQPSK for digital cellular radio," *IEEE Trans. Veh. Tech*., pp. 46–57, Febmary 1993.

[22] P. A. Bello and B. D. Nelin, "The effects of frequency selective fading on the binary error probabilities of incoherent and differentially coherent matched filter receivers," *IEEE Trans. Commun. Syst*., pp. 170–86, June 1963.

[23] M. B. Pursley, *Introduction to Digital Communications*, Prentice-Hall, Englewood Cliffs, NJ, 2005.

[24] J. C.-I. Chuang, "The effects of time delay spread on portable radio communications channels with digital modulation," *IEEE J. Sel. Areas Commun*., pp. 879–89, June 1987.

[25] N. R. Sollenberger, J. C. I. Chuang, L. F. Chang, S. Ariyavisitakul and H. Arnold, "Architecture and implementation of an efficient and robust TDMA frame structure for digital portable communications," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 1, pp. 250-260, Feb. 1991.

[26] J. C. I. Chuang, "The effects of delay spread on 2-PSK, 4-PSK, 8-PSK and 16-QAM in a portable radio environment," *IEEE Transactions on Vehicular Technology*, vol. 38, no. 2, pp. 43-45, May 1989.

[27] C. Liu and K. Feher, "Bit error rate performance of $\pi/4$ DQPSK in a frequency selective fast Rayleigh fading channel," *IEEE Trans. Veh. Tech*., pp. 558–68, August 1991.

[28] S. Gumnathan and K. Feher, "Multipath simulation models for mobile radio channels," *Proc. IEEE Veh. Tech. Conf*., pp. 131–4, May 1992.

[29] A. F. Molisch, *Wireless Communications: From Fundamentals to Beyond 5G* , 3rd Ed, Wiley-IEEE Press, 2022.

Chapter 7

Diversity

In Chapter 6 we saw that both Rayleigh fading and log-normal shadowing exact a large power penalty on the performance of modulation over wireless channels. One of the best techniques to mitigate the effects of fading is diversity combining of independently fading signal paths. Diversity combining exploits the fact that independent signal paths have a low probability of experiencing deep fades simultaneously. Thus, the idea behind diversity is to send the same data over independent fading paths. These independent paths are combined in such a way that the fading of the resultant signal is reduced. For example, consider a system with two antennas at either the transmitter or receiver that experience independent fading. If the antennas are spaced sufficiently far apart, it is unlikely that they both experience deep fades at the same time. By selecting the antenna with the strongest signal, a technique known as *selection combining*, we obtain a much better signal than if we had just one antenna. This chapter focuses on common methods used at the transmitter and receiver to achieve diversity. A more detailed treatment of diversity combining methods and their performance, including the effects of correlated fading on the different paths as well as channel estimation errors, can be found in [1, Chap. 9].

Diversity techniques that mitigate the effect of multipath fading are called *microdiversity*, and that is the focus of this chapter. Diversity to mitigate the effects of shadowing from buildings and objects is called *macrodiversity*. Macrodiversity is generally implemented by combining signals received by several base stations or access points, which requires coordination among these different stations or points. Such coordination is implemented as part of the networking protocols in infrastructure-based wireless networks. We will therefore defer discussion of macrodiversity until Chapter 15, where we examine the design of such networks.

7.1 Realization of Independent Fading Paths

There are many ways of achieving independent fading paths in a wireless system. One method is to use multiple transmit or receive antennas, also called an antenna array, where the elements of the array are separated in distance. This type of diversity is referred to as *space diversity*. Note that with receiver space diversity, independent fading paths are realized without an increase in transmit signal power or bandwidth. Moreover, coherent combining of the diversity signals increases the signal-to-noise power ratio at the receiver over the SNR that would be obtained with just a single receive antenna. This SNR increase, called array gain, can also be obtained with transmitter space diversity by appropriately weighting the antenna transmit powers relative to the channel gains. In addition to array gain, space diversity also provides diversity gain, defined as the change in slope of the error probability resulting from the diversity combining. We will describe both the array gain and diversity gain for specific diversity-combining techniques in subsequent sections.

The maximum diversity gain for either transmitter or receiver space diversity typically requires that the separation between antennas be such that the fading amplitudes corresponding to each antenna are approximately

independent. For example, from equation (3.30) we know that, in a uniform scattering environment with omni-directional transmit and receive antennas, the minimum antenna separation required for independent fading on each antenna is approximately one half-wavelength ($.38\lambda$, to be exact). If the transmit or receive antennas are directional, the scattering is not uniform, or a dominant LOS component exists, then a larger antenna separation is typically required to obtain independent fading samples [2, Chap. 12].

A second method of achieving diversity is by using either two transmit antennas or two receive antennas with different polarization (e.g., vertically and horizontally polarized waves). The two transmitted waves follow the same path. However, since the multiple random reflections distribute the power nearly equally relative to both polarizations, the average receive power corresponding to either polarized antenna is approximately the same. Since the scattering angle relative to each polarization is random, it is highly improbable that signals received on the two differently polarized antennas would be simultaneously in a deep fade. There are two disadvantages of polarization diversity. First, you can have at most two diversity branches, corresponding to the two types of polarization. The second disadvantage is that polarization diversity loses effectively half the power (3 dB) because the transmit or receive power is divided between the two differently polarized antennas.

Directional antennas provide angle (or directional) diversity by restricting the receive antenna beamwidth to a given angle. In the extreme, if the angle is very small then at most one of the multipath rays will fall within the receive beamwidth, so there is no multipath fading from multiple rays. However, this diversity technique requires either a sufficient number of directional antennas to span all possible directions of arrival or a single antenna whose directivity can be steered to the arrival angle of one of the multipath components (preferably the strongest one). Note also that with this technique the SNR may decrease owing to the loss of multipath components that fall outside the receive antenna beamwidth – unless the directional gain of the antenna is sufficiently large to compensate for this lost power. *Smart antennas* are antenna arrays with adjustable phase at each antenna element: such arrays form directional antennas that can be steered to the incoming angle of the strongest multipath component [3].

Frequency diversity is achieved by transmitting the same narrowband signal at different carrier frequencies, where the carriers are separated by the coherence bandwidth of the channel. This technique requires additional transmit power to send the signal over multiple frequency bands. Spread-spectrum techniques, discussed in Chapter 13, are sometimes described as providing frequency diversity in that the channel gain varies across the bandwidth of the transmitted signal. However, this is not equivalent to sending the same information signal over independently fading paths. As discussed in Section 13.2.4, spread spectrum with RAKE reception does provide independently fading paths of the information signal and thus is a form of path diversity. Time diversity is achieved by transmitting the same signal at different times, where the time difference is greater than the channel coherence time (the inverse of the channel Doppler spread). Time diversity does not require increased transmit power but it does lower the data rate, since data is repeated in the diversity time slots rather than sending new data in those time slots. Time diversity can also be achieved through coding and interleaving, as will be discussed in Chapter 8. Clearly time diversity cannot be used for stationary applications, since the channel coherence time is infinite and thus fading is highly correlated over time.

In this chapter we focus on space diversity as a reference when describing diversity systems and the different combining techniques, although the techniques can be applied to any type of diversity. Thus, the combining techniques will be defined as operations on an antenna array. Receiver and transmitter diversity are treated separately, since their respective system models and combining techniques have important differences.

7.2 Receiver Diversity

7.2.1 System Model

In receiver diversity, the independent fading paths associated with multiple receive antennas are combined to obtain a signal that is then passed through a standard demodulator. The combining can be done in several ways, which vary in complexity and overall performance. Most combining techniques are linear: the output of the combiner is just a weighted sum of the different fading paths or *branches*, as shown in Figure 7.1 for M-branch diversity. In this model the complex fading associated with the ith branch is $r_i e^{j\theta_i}$. The signal on the ith branch is multiplied by complex gain α_i and then summed with the other weighted branches to form the combiner output. When all but one of the complex α_i are zero, only one path is passed to the combiner output. If more than one of the α_i are nonzero then the combiner adds together multiple paths, though each path may be weighted by a different value. Combining more than one branch signal requires *co-phasing*, where the phase θ_i of the ith branch is removed through multiplication by $\alpha_i = a_i e^{-j\theta_i}$ for some real-valued a_i. Without co-phasing, the branch signals would not add up coherently in the combiner, so the resulting output could still exhibit significant fading due to constructive and destructive addition of the signals in all the branches.

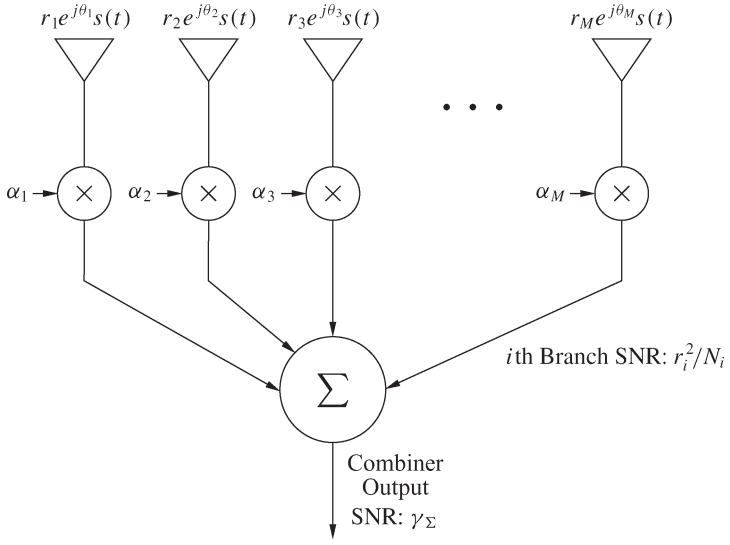


Figure 7.1: Linear combiner.

Since co-phasing requires knowledge of the branch phase prior to multiplication by α_i, coherent detection to determine this phase value is required in each branch, as shown in Figure 7.2. This coherent detection can also determine the value of the branch signal envelope r_i or SNR $\gamma_i = r_i^2/N_i$ for N_i the noise power on the ith branch,[1] which α_i depends on in some combining techniques. A dedicated coherent detector for each branch of a diversity system significantly increases the hardware complexity and power consumption relative to a single-branch system, particularly for a large number of branches. When the combiner utilizes just one branch at any given time, i.e. $\alpha_i = 0$ for all except one of the M branches, then the coherent detector can be placed after the combiner in Figure 7.1, as co-phasing multiple signals is no longer needed. We will assume throughout this chapter that the noise power in each branch $N_i = N = N_0 B$ is the same. When that is not the case, the results of this chapter can be applied by scaling the amplitude term r_i by $\sqrt{N/N_i}$ for average noise power $N = N_0 B$ across all branches.

The main purpose of diversity is to coherently combine the independent fading paths so that the effects of

[1] In practice $r_i^2 + N_i$ is easier to measure than SNR, since the former involves finding only the total power in the received signal.

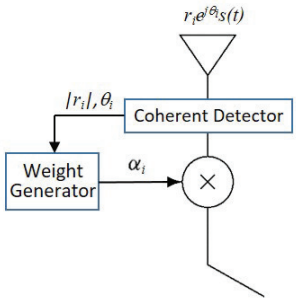


Figure 7.2: Branch Coherent Detection

fading are mitigated. The signal output from the combiner equals the original transmitted signal $s(t)$ multiplied by a random real-valued amplitude term $\alpha_\Sigma = \sum_i a_i r_i$. This amplitude term results in a random SNR γ_Σ at the combiner output, where the distribution of γ_Σ is a function of the number of diversity paths, the fading distribution on each path, and the combining technique, as we shall describe in more detail.

The array gain in receiver space diversity results from coherent combining of multiple receive signals. Even in the absence of fading, this can lead to an increase in average received SNR. For example, suppose there is no fading and that each branch has the same SNR $\gamma_i = r_i^2/(N_0 B) = \gamma$. Since both the noise and the signal are multiplied by a_i, the SNR of the combined output is

$$\gamma_\Sigma = \frac{\left(\sum_{i=1}^M a_i r_i\right)^2}{\sum_{i=1}^M a_i^2 N_0 B}. \tag{7.1}$$

Now set $a_i = r_i$ in (7.1). We will see later that setting $a_i = cr_i$ for any constant c is optimal for maximal-ratio combining in fading. Then,

$$\gamma_\Sigma = \frac{\left(\sum_{i=1}^M r_i^2\right)^2}{N_0 B \sum_{i=1}^M r_i^2} = \frac{\sum_{i=1}^M r_i^2}{N_0 B} = M\gamma. \tag{7.2}$$

Thus, in the absence of fading, with appropriate weighting there is an M-fold increase in SNR due to the coherent combining of the M signals received from the different antennas. This SNR increase in the absence of fading is referred to as the *array gain*. More precisely, array gain A_g is defined as the increase in the average combined SNR $\bar{\gamma}_\Sigma$ over the average branch SNR $\bar{\gamma}$:

$$A_g = \frac{\bar{\gamma}_\Sigma}{\bar{\gamma}}. \tag{7.3}$$

Array gain occurs for all diversity combining techniques but is most pronounced in maximal-ratio combining. The array gain allows a system with multiple transmit or receive antennas in a fading channel to achieve better performance than a system without diversity in an AWGN channel with the same average SNR. We will see this effect in performance curves for both maximal-ratio and equal-gain combining with a large number of antennas.

With fading, the combining of multiple independent fading paths leads to a more favorable distribution for γ_Σ than would be the case with just a single path. In particular, the performance of a diversity system (whether it uses space diversity or another form of diversity) in terms of $\bar{P}_s$ and P_{out} is as defined in Sections 6.3.1 and 6.3.2:

$$\bar{P}_s = \int_0^\infty P_s(\gamma) p_{\gamma_\Sigma}(\gamma) d\gamma, \tag{7.4}$$

where $P_s(\gamma)$ is the probability of symbol error for demodulation of $s(t)$ in AWGN with SNR γ; and

$$P_{\text{out}} = p(\gamma_\Sigma \leq \gamma_0) = \int_0^{\gamma_0} p_{\gamma_\Sigma}(\gamma) d\gamma \tag{7.5}$$

for some target SNR value γ_0. The more favorable distribution for γ_Σ resulting from diversity combining leads to a decrease in $\bar{P}_s$ and P_{out}, and the resulting performance advantage is called the *diversity gain*.

When $\bar{P}_s$ as a function of $p_{\gamma_\Sigma}(\gamma)$ has a particular form, the performance benefit of diversity combining is defined in terms of the *diversity order* of the system. In particular, if the average probability of error with diversity combining can be expressed in the form $\bar{P}_s = c\bar{\gamma}^{-M}$, where c is a constant that depends on the specific modulation and coding and $\bar{\gamma}$ is the average received SNR per branch, then M is defined as the *diversity order* of the system. The diversity order indicates how the *slope* of the average probability of error as a function of average SNR changes with diversity (Figures 7.4 and 7.7 below show these slope changes as a function of M for different combining techniques). Recall from (6.72) that a general approximation for average error probability in Rayleigh fading with no diversity is $\bar{P}_s \approx \alpha_M/(2\beta_M\bar{\gamma})$. This expression has a diversity order of 1, which is consistent with a single fading path, e.g. from a single receive antenna. The maximum diversity order of a system with M fading paths, e.g. from M antennas, is M. When the diversity order equals the number of independent fading paths that are combined via diversity, the system is said to achieve *full diversity order*.

In the following sections we will describe the different combining techniques and their performance in more detail. These techniques entail various trade-offs between performance and complexity.

7.2.2 Selection Combining

In selection combining (SC), the combiner outputs the signal on the branch with the highest SNR. Because only one branch is used at a time, SC requires just one receiver that is switched into the active antenna branch. However, a dedicated receiver on each antenna branch may be needed for systems that transmit continuously in order to monitor SNR on each branch simultaneously. With SC, the path output from the combiner has an SNR equal to the maximum SNR of all the branches. Moreover, since only one branch output is used, co-phasing of multiple branches is not required; hence this technique can be used with either coherent or differential modulation.

For M-branch diversity, the cumulative distribution function (cdf) of γ_Σ is given by

$$P_{\gamma_\Sigma}(\gamma) = p(\gamma_\Sigma < \gamma) = p(\max[\gamma_1, \gamma_2, \ldots, \gamma_M] < \gamma) = \prod_{i=1}^{M} p(\gamma_i < \gamma). \tag{7.6}$$

We obtain the distribution of γ_Σ by differentiating $P_{\gamma_\Sigma}(\gamma)$ relative to γ, and we obtain the outage probability by evaluating $P_{\gamma_\Sigma}(\gamma)$ at $\gamma = \gamma_0$. Assume that we have M branches with uncorrelated Rayleigh fading amplitudes r_i and corresponding instantaneous SNR $\gamma_i = r_i^2/N_0B$. Defining the average SNR on the ith branch as $\bar{\gamma}_i = \mathbf{E}[\gamma_i]$, the SNR distribution will be exponential:

$$p(\gamma_i) = \frac{1}{\bar{\gamma}_i} e^{-\gamma_i/\bar{\gamma}_i}. \tag{7.7}$$

From (6.57), the outage probability for a target γ_0 on the ith branch in Rayleigh fading is

$$P_{\text{out}}(\gamma_0) = 1 - e^{-\gamma_0/\bar{\gamma}_i}. \tag{7.8}$$

The outage probability of the selection combiner for the target γ_0 is then

$$P_{\text{out}}(\gamma_0) = \prod_{i=1}^{M} p(\gamma_i < \gamma_0) = \prod_{i=1}^{M} [1 - e^{-\gamma_0/\bar{\gamma}_i}]. \tag{7.9}$$

If the average SNR for all of the branches are the same ($\bar{\gamma}_i = \bar{\gamma}$ for all i), then this reduces to

$$P_{\text{out}}(\gamma_0) = p(\gamma_\Sigma < \gamma_0) = [1 - e^{-\gamma_0/\bar{\gamma}}]^M. \tag{7.10}$$

Differentiating (7.10) relative to γ_0 yields the distribution for γ_Σ:

$$p_{\gamma_\Sigma}(\gamma) = \frac{M}{\bar{\gamma}}[1 - e^{-\gamma/\bar{\gamma}}]^{M-1} e^{-\gamma/\bar{\gamma}}. \tag{7.11}$$

From (7.11) we see that the average SNR of the combiner output in independent and identically distributed Rayleigh fading is

$$\begin{aligned} \bar{\gamma}_\Sigma &= \int_0^\infty \gamma p_{\gamma_\Sigma}(\gamma) d\gamma \\ &= \int_0^\infty \frac{\gamma M}{\bar{\gamma}}[1 - e^{-\gamma/\bar{\gamma}}]^{M-1} e^{-\gamma/\bar{\gamma}} d\gamma \\ &= \bar{\gamma} \sum_{i=1}^{M} \frac{1}{i}, \end{aligned} \tag{7.12}$$

where (7.12) follows from the binomial expansion of $[1 - e^{-\gamma/\bar{\gamma}}]^{M-1}$ and induction to obtain the final result. Thus, the average SNR gain and corresponding array gain increase with M, but not linearly. The biggest gain is obtained by going from no diversity to two-branch diversity. Increasing the number of diversity branches from two to three will give much less gain than going from one to two, and in general increasing M yields diminishing returns in terms of the array gain. This trend is also illustrated in Figure 7.3, which shows P_{out} versus $\bar{\gamma}/\gamma_0$ for different M in i.i.d. Rayleigh fading. We see that there is dramatic improvement even with two-branch selection combining: going from $M = 1$ to $M = 2$ at 1% outage probability yields an approximate 12-dB reduction in required SNR, and at .01% outage probability there is an approximate 20-dB reduction in required SNR. However, at .01% outage, going from two-branch to three-branch diversity results in an additional reduction of about 7 dB, and going from three-branch to four-branch results in an additional reduction of about 4 dB. Clearly the power savings is most substantial when going from no diversity to two-branch diversity, with diminishing returns as the number of branches is increased. It should be noted also that, even with Rayleigh fading on all branches, the distribution of the combiner output SNR is no longer exponential.

Example 7.1: Find the outage probability of BPSK modulation at $P_b = 10^{-3}$ for a Rayleigh fading channel with SC diversity for $M = 1$ (no diversity), $M = 2$, and $M = 3$. Assume equal branch SNRs of $\bar{\gamma} = 15\,\text{dB}$.
Solution: A BPSK modulated signal with $\gamma_b = 7\,\text{dB}$ has $P_b = 10^{-3}$. Thus, we have $\gamma_0 = 7\,\text{dB}$. Substituting $\gamma_0 = 10^{.7}$ and $\bar{\gamma} = 10^{1.5}$ into (7.10) yields $P_{\text{out}} = .1466$ for $M = 1$, $P_{\text{out}} = .0215$ for $M = 2$, and $P_{\text{out}} = .0031$ for $M = 3$. We see that each additional branch reduces outage probability by almost an order of magnitude.

The average probability of symbol error is obtained from (7.4) with $P_s(\gamma)$ the probability of symbol error in AWGN for the signal modulation and $p_{\gamma_\Sigma}(\gamma)$ the distribution of the combiner SNR. For most fading distributions and coherent modulations, this result cannot be obtained in closed form and must be evaluated numerically or by approximation. In Figure 7.4 we plot $\bar{P}_b$ versus $\bar{\gamma}_b$ in i.i.d. Rayleigh fading, obtained by a numerical evaluation of $\int Q(\sqrt{2\gamma}) p_{\gamma_\Sigma}(\gamma) d\gamma$ for $p_{\gamma_\Sigma}(\gamma)$ given by (7.11). The figure shows that the diversity system for $M \geq 8$ has a lower error probability than an AWGN channel with the same SNR, owing to the array gain of the combiner. The same will be true for the performance of maximal-ratio and equal-gain combining. Closed-form results do exist for differential modulation under i.i.d. Rayleigh fading on each branch (see [1, Chap. 9.7] and [4, Chap. 6.1]). For

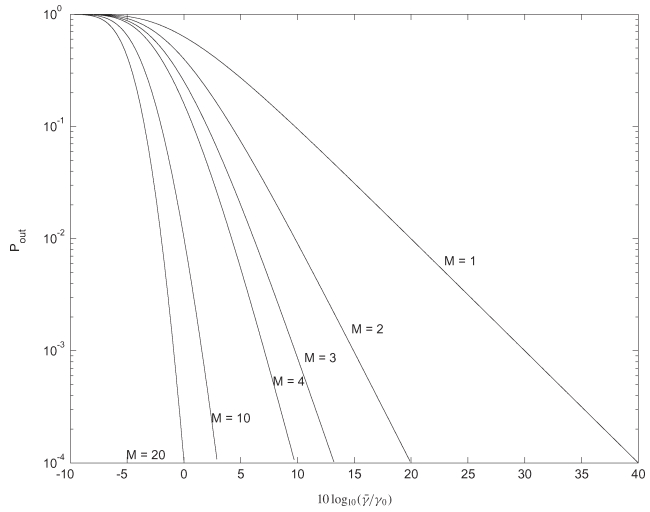


Figure 7.3: Outage probability of selection combining in Rayleigh fading.

example, it can be shown for DPSK with $p_{\gamma_\Sigma}(\gamma)$ given by (7.11) that the average probability of symbol error is given by

$$\bar{P}_b = \int_0^\infty \frac{1}{2} e^{-\gamma} p_{\gamma_\Sigma}(\gamma) d\gamma = \frac{M}{2} \sum_{m=0}^{M-1} (-1)^m \frac{\binom{M-1}{m}}{1+m+\bar{\gamma}}. \tag{7.13}$$

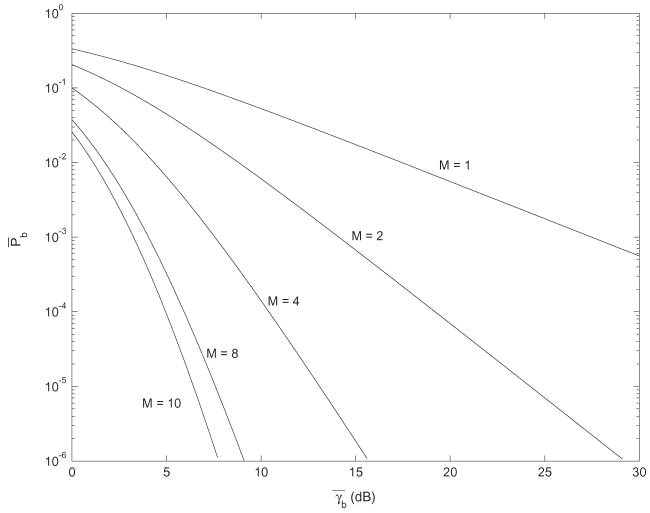


Figure 7.4: Average P_b of BPSK under selection combining with i.i.d. Rayleigh fading.

In the foregoing derivations we assume that there is no correlation between the branch amplitudes. If the correlation is nonzero then there is a slight degradation in performance, which is almost negligible for correlations below .5. Derivation of the exact performance degradation due to branch correlation can be found in [1, Chap. 9.7].

7.2.3 Threshold Combining

Selection combining for systems that transmit continuously may require a dedicated receiver on each branch to continuously monitor branch SNR. A simpler type of combining, called threshold combining, avoids the need for a dedicated receiver on each branch by scanning each of the branches in sequential order and outputting the first signal whose SNR is above a given threshold γ_T. As in SC, co-phasing is not required because only one branch output is used at a time. Hence this technique can be used with either coherent or differential modulation.

Once a branch is chosen, the combiner outputs that signal as long as the SNR on that branch remains above the desired threshold. If the SNR on the selected branch falls below the threshold, the combiner switches to another branch. There are several criteria the combiner can use for determining which branch to switch to [5]. The simplest criterion is to switch randomly to another branch. With only two-branch diversity this is equivalent to switching to the other branch when the SNR on the active branch falls below γ_T. This method is called *switch-and-stay combining* (SSC). The switching process and SNR associated with SSC is illustrated in Figure 7.5. Since the SSC does not select the branch with the highest SNR, as illustrated in Figure 7.5 on the third switch, its performance is between that of no diversity and ideal SC.

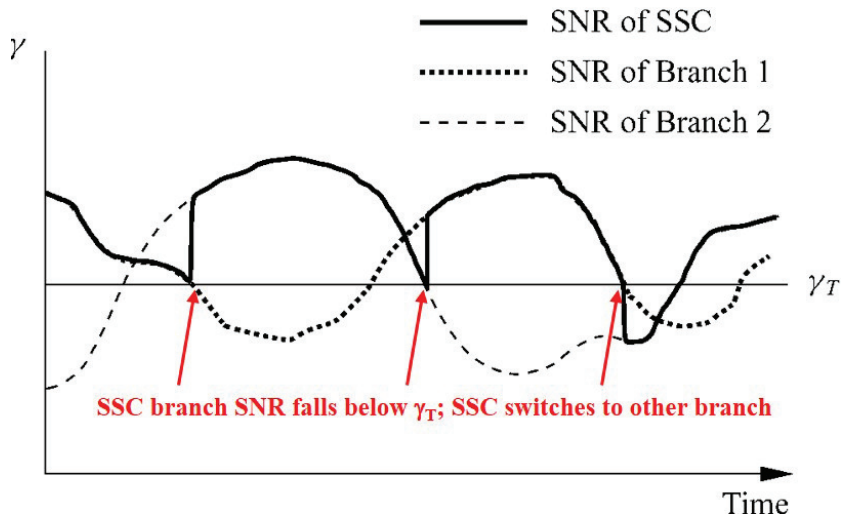


Figure 7.5: SNR of the SSC technique. Switching when the SNR falls below a threshold does not always select the branch with the highest SNR.

Let us denote the SNR on the ith branch by γ_i and the SNR of the combiner output by γ_Σ. The cdf of γ_Σ will depend on the threshold level γ_T and the cdf of γ_i. For two-branch diversity with i.i.d. branch statistics, the cdf of the combiner output $P_{\gamma_\Sigma}(\gamma) = p(\gamma_\Sigma \leq \gamma)$ can be expressed in terms of the cdf $P_{\gamma_i}(\gamma) = p(\gamma_i \leq \gamma)$ and the distribution $p_{\gamma_i}(\gamma)$ of the individual branch SNRs as

$$P_{\gamma_\Sigma}(\gamma) = \begin{cases} P_{\gamma_1}(\gamma_T)P_{\gamma_2}(\gamma) & \gamma < \gamma_T, \\ p(\gamma_T \leq \gamma_2 \leq \gamma) + P_{\gamma_1}(\gamma_T)P_{\gamma_2}(\gamma) & \gamma \geq \gamma_T. \end{cases} \tag{7.14}$$

For Rayleigh fading in each branch with $\bar{\gamma}_i = \bar{\gamma}$ $(i = 1, 2)$, this yields

$$P_{\gamma_\Sigma}(\gamma) = \begin{cases} 1 - e^{-\gamma_T/\bar{\gamma}} - e^{-\gamma/\bar{\gamma}} + e^{-(\gamma_T+\gamma)/\bar{\gamma}} & \gamma < \gamma_T, \\ 1 - 2e^{-\gamma/\bar{\gamma}} + e^{-(\gamma_T+\gamma)/\bar{\gamma}} & \gamma \geq \gamma_T. \end{cases} \tag{7.15}$$

The outage probability P_{out} associated with a given γ_0 is obtained by evaluating $P_{\gamma_\Sigma}(\gamma)$ at $\gamma = \gamma_0$:

$$P_{\text{out}}(\gamma_0) = P_{\gamma_\Sigma}(\gamma_0) = \begin{cases} 1 - e^{-\gamma_T/\bar{\gamma}} - e^{-\gamma_0/\bar{\gamma}} + e^{-(\gamma_T+\gamma_0)/\bar{\gamma}} & \gamma_0 < \gamma_T, \\ 1 - 2e^{-\gamma_0/\bar{\gamma}} + e^{-(\gamma_T+\gamma_0)/\bar{\gamma}} & \gamma_0 \geq \gamma_T. \end{cases} \tag{7.16}$$

The performance of SSC under other types of fading, as well as the effects of fading correlation, is studied in [1, Chap. 9.8] and [6, 7]. In particular, it is shown in [1, Chap. 9.8] that, for any fading distribution, SSC with an optimized threshold of $\gamma_T = \gamma_0$ has the same outage probability as SC.

Example 7.2: Find the outage probability of BPSK modulation at $P_b = 10^{-3}$ for two-branch SSC diversity with i.i.d. Rayleigh fading on each branch for threshold values of $\gamma_T = 5\,\text{dB}$, 7 dB, and 10 dB. Assume the average branch SNR is $\bar{\gamma} = 15\,\text{dB}$. Discuss how the outage probability changes with γ_T. Also compare outage probability under SSC with that of SC and no diversity from Example 7.1.

Solution: As in Example 7.1, $\gamma_0 = 7\,\text{dB}$. For $\gamma_T = 5\,\text{dB}$ we have $\gamma_0 \geq \gamma_T$, so we use the second line of (7.16) to get

$$P_{\text{out}} = 1 - 2e^{-10^{.7}/10^{1.5}} + e^{-(10^{.5}+10^{.7})/10^{1.5}} = .0654. \tag{7.17}$$

For $\gamma_T = 7\,\text{dB}$ we have $\gamma_0 = \gamma_T$, so we again use the second line of (7.16) and obtain

$$P_{\text{out}} = 1 - 2e^{-10^{.7}/10^{1.5}} + e^{-(10^{.7}+10^{.7})/10^{1.5}} = .0215. \tag{7.18}$$

For $\gamma_T = 10\,\text{dB}$ we have $\gamma_0 < \gamma_T$, so we use the first line of (7.16) to get

$$P_{\text{out}} = 1 - e^{-10/10^{1.5}} - e^{-10^{.7}/10^{1.5}} + e^{-(10+10^{.7})/10^{1.5}} = .0397. \tag{7.19}$$

We see that the outage probability is smaller for $\gamma_T = 7\,\text{dB}$ than for the other two values. At $\gamma_T = 5\,\text{dB}$ the threshold is too low, so the active branch can be below the target γ_0 for a long time before a switch is made; this contributes to a large outage probability. At $\gamma_T = 10\,\text{dB}$ the threshold is too high: the active branch will often fall below this threshold value, which will cause the combiner to switch to the other antenna even though that other antenna may have a lower SNR than the active one. This example shows that the threshold γ_T minimizing P_{out} equals γ_0.

From Example 7.1, SC has $P_{\text{out}} = .0215$. Thus, $\gamma_T = 7\,\text{dB}$ is the optimal threshold where SSC performs the same as SC. We also see that performance with an unoptimized threshold can be much worse than SC. However, the performance of SSC under all three thresholds is better than the performance without diversity, derived as $P_{\text{out}} = .1466$ in Example 7.1.

We obtain the distribution of γ_Σ by differentiating (7.14) relative to γ. Then the average probability of error is obtained from (7.4) with $P_s(\gamma)$ the probability of symbol error in AWGN and $p_{\gamma_\Sigma}(\gamma)$ the distribution of the SSC output SNR. For most fading distributions and coherent modulations, this result cannot be obtained in closed form and so must be evaluated numerically or by approximation. However, for i.i.d. Rayleigh fading we can differentiate (7.15) to obtain

$$p_{\gamma_\Sigma}(\gamma) = \begin{cases} (1 - e^{-\gamma_T/\bar{\gamma}})(1/\bar{\gamma})e^{-\gamma/\bar{\gamma}} & \gamma < \gamma_T, \\ (2 - e^{-\gamma_T/\bar{\gamma}})(1/\bar{\gamma})e^{-\gamma/\bar{\gamma}} & \gamma \geq \gamma_T. \end{cases} \tag{7.20}$$

As with SC, for most fading distributions and coherent modulations the resulting average probability of error is not in closed form and must be evaluated numerically. However, closed form results do exist for differential

modulation under i.i.d. Rayleigh fading on each branch. In particular, the average probability of symbol error for DPSK is given by

$$\bar{P}_b = \int_0^\infty \frac{1}{2} e^{-\gamma} p_{\gamma_\Sigma}(\gamma)\, d\gamma = \frac{1}{2(1+\bar{\gamma})}(1 - e^{-\gamma_T/\bar{\gamma}} + e^{-\gamma_T} e^{-\gamma_T/\bar{\gamma}}). \tag{7.21}$$

Example 7.3: Find the average probability of error for DPSK modulation under two-branch SSC diversity with i.i.d. Rayleigh fading on each branch for threshold values of $\gamma_T = 3\,\text{dB}$, 7 dB, and 10 dB. Assume the average branch SNR is $\bar{\gamma} = 15\,\text{dB}$. Discuss how the average probability of error changes with γ_T. Also compare average error probability under SSC with that of SC and with no diversity.

Solution: Evaluating (7.21) with $\bar{\gamma} = 15\,\text{dB}$ and $\gamma_T = 3, 7$, and 10 dB yields (respectively) $\bar{P}_b = .0029$, $P_b = .0023$, and $\bar{P}_b = .0042$. As in the previous example, there is an optimal threshold that minimizes average probability of error. Setting the threshold too high or too low degrades performance. From (7.13) we have that SC yields $\bar{P}_b = .5(1+10^{1.5})^{-1} - .5(2+10^{1.5})^{-1} = 4.56 \cdot 10^{-4}$, which is roughly an order of magnitude less than with SSC and an optimized threshold. With no diversity we have $\bar{P}_b = .5(1+10^{1.5})^{-1} = .0153$, which is roughly an order of magnitude worse than with two-branch SSC.

7.2.4 Equal-Gain Combining

Equal-gain combining (EGC) is a more complex technique than those discussed in prior sections, as it co-phases the signals on each branch and then combines them with equal weighting, $\alpha_i = e^{-\theta_i}$. The co-phasing requires a coherent phase detector on each branch, but the instantaneous SNR of the branch is not required, in contrast to maximal-ratio combining. The SNR of the combiner output, assuming equal noise power $N_0 B$ in each branch, is then given by

$$\gamma_\Sigma = \frac{1}{MN_0B}\left(\sum_{i=1}^{M} r_i\right)^2. \tag{7.22}$$

The distribution and cdf of γ_Σ do not exist in closed form. For i.i.d. Rayleigh fading with two-branch diversity and average branch SNR $\bar{\gamma}$, an expression for the cdf in terms of the Q-function can be derived (see [4, Chap. 6.4] and [8, Chap. 5.6]) as

$$P_{\gamma_\Sigma}(\gamma) = 1 - e^{-2\gamma/\bar{\gamma}} - \sqrt{\pi\gamma/\bar{\gamma}}\, e^{-\gamma/\bar{\gamma}}(1 - 2Q(\sqrt{2\gamma/\bar{\gamma}})). \tag{7.23}$$

The resulting outage probability is given by

$$P_{\text{out}}(\gamma_0) = 1 - e^{-2\gamma_R} - \sqrt{\pi\gamma_R}\, e^{-\gamma_R}(1 - 2Q(\sqrt{2\gamma_R})), \tag{7.24}$$

where $\gamma_R = \gamma_0/\bar{\gamma}$. Differentiating (7.23) with respect to γ yields the distribution

$$p_{\gamma_\Sigma}(\gamma) = \frac{1}{\bar{\gamma}} e^{-2\gamma/\bar{\gamma}} + \sqrt{\pi} e^{-\gamma/\bar{\gamma}} \left(\frac{1}{\sqrt{4\gamma\bar{\gamma}}} - \frac{1}{\bar{\gamma}}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right)\left(1 - 2Q\left(\sqrt{\frac{2\gamma}{\bar{\gamma}}}\right)\right). \tag{7.25}$$

Substituting this into (7.4) for BPSK yields the average probability of bit error:

$$\bar{P}_b = \int_0^\infty Q(\sqrt{2\gamma}) p_{\gamma_\Sigma}(\gamma) d\gamma = .5\left(1 - \sqrt{1 - \left(\frac{1}{1+\bar{\gamma}}\right)^2}\right). \tag{7.26}$$

7.2.5 Maximal-Ratio Combining

In maximal-ratio combining (MRC) the output is a weighted sum of all branches, so the α_i in Figure 7.1 are all nonzero, and the weights are determined to maximize the combiner output's SNR. The signals are co-phased and so $\alpha_i = a_i e^{-j\theta_i}$, where θ_i is the phase of the incoming signal on the ith branch. Thus, the envelope of the combiner output will be $r = \sum_{i=1}^{M} a_i r_i$. Assuming the same noise power $N_0 B$ in each branch yields a total noise power at the combiner output of $\sum_{i=1}^{M} a_i^2 N_0 B$. Thus, the output SNR of the combiner is

$$\gamma_\Sigma = \frac{1}{N_0 B} \frac{\left(\sum_{i=1}^{M} a_i r_i\right)^2}{\sum_{i=1}^{M} a_i^2}. \tag{7.27}$$

The goal of MRC is to choose the a_i to maximize γ_Σ. Intuitively, branches with a high SNR should be weighted more than branches with a low SNR. We find the a_is that maximize γ_Σ by taking partial derivatives of (7.27) or finding the weights that achieve equality in the Cauchy-Schwartz inequality[2]. Solving for the optimal weights yields $a_i^2 = c\gamma_i = c r_i^2/(N_0 B)$ for any constant c where the constant c cancels out in the SNR calculation. Setting $a_i = r_i$ yields the combiner SNR

$$\gamma_\Sigma = \sum_{i=1}^{M} \frac{r_i^2}{N_0 B} = \sum_{i=1}^{M} \gamma_i. \tag{7.28}$$

Thus, the SNR of the combiner output is the sum of SNRs on each branch. Hence, the average combiner SNR and corresponding array gain increase linearly with the number of diversity branches M, in contrast to the diminishing returns associated with the average combiner SNR in SC given by (7.12). As with SC, the distribution of the combiner output SNR does not remain exponential even when there is Rayleigh fading on all branches.

To obtain the distribution of γ_Σ, we take the product of the exponential moment generating functions or characteristic functions. Assuming i.i.d. Rayleigh fading on each branch with equal average branch SNR $\bar{\gamma}$, the distribution of γ_Σ is χ^2 with $2M$ degrees of freedom, expected value $\bar{\gamma}_\Sigma = M\bar{\gamma}$, and variance $2M\bar{\gamma}$:

$$p_{\gamma_\Sigma}(\gamma) = \frac{\gamma^{M-1} e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}^M (M-1)!}, \quad \gamma \geq 0. \tag{7.29}$$

The corresponding outage probability for a given threshold γ_0 is given by

$$P_{\text{out}} = p(\gamma_\Sigma < \gamma_0) = \int_0^{\gamma_0} p_{\gamma_\Sigma}(\gamma)\, d\gamma = 1 - e^{-\gamma_0/\bar{\gamma}} \sum_{k=1}^{M} \frac{(\gamma_0/\bar{\gamma})^{k-1}}{(k-1)!}. \tag{7.30}$$

Figure 7.6 is a plot of P_{out} for maximal-ratio combining indexed by the number of diversity branches.

The average probability of symbol error is obtained from (7.4) with $P_s(\gamma)$ the probability of symbol error in AWGN for the signal modulation and $p_{\gamma_\Sigma}(\gamma)$ the distribution of γ_Σ. For BPSK modulation with i.i.d. Rayleigh fading, where $p_{\gamma_\Sigma}(\gamma)$ is given by (7.29), it can be shown (see [4, Chap. 6.3]) that

$$\bar{P}_b = \int_0^\infty Q(\sqrt{2\gamma}) p_{\gamma_\Sigma}(\gamma)\, d\gamma = \left(\frac{1-\Gamma}{2}\right)^M \sum_{m=0}^{M-1} \binom{M-1+m}{m} \left(\frac{1+\Gamma}{2}\right)^m, \tag{7.31}$$

where $\Gamma = \sqrt{\bar{\gamma}/(1+\bar{\gamma})}$. This equation is plotted in Figure 7.7. Comparing the outage probability for MRC in Figure 7.6 with that of SC in Figure 7.3, or comparing the average probability of error for MRC in Figure 7.7 with

[2]The Cauchy-Schwartz inequality for two vectors $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ is $\sum_i x_i y_i \leq \sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}$

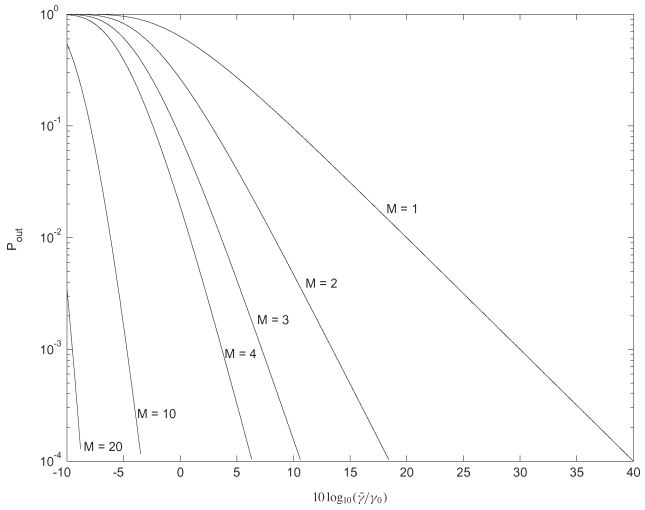


Figure 7.6: P_{out} for maximal-ratio combining with i.i.d. Rayleigh fading.

that of SC in Figure 7.4, indicates that MRC has significantly better performance than SC. In Section 7.4 we will use a different analysis based on moment generating functions to compute average error probability under MRC, which can be applied to any modulation type, any number of diversity branches, and any fading distribution on the different branches.

For modulations whose probability of error in AWGN can be approximated as $P_s = \alpha Q(\sqrt{\beta\gamma})$ for SNR γ, e.g., BPSK, MPSK, and MQAM, we can obtain a simple upper bound on the average probability of error by applying the Chernoff bound $Q(x) \leq e^{-x^2/2}$ from (5.46) to the Q-function. For these modulations, we have

$$P_s = \alpha Q(\sqrt{\beta\gamma_\Sigma}) \leq \alpha e^{-\beta\gamma_\Sigma/2} = \alpha e^{-\beta(\gamma_1+\cdots+\gamma_M)/2}. \tag{7.32}$$

Integrating over the χ^2 distribution for γ_Σ yields

$$\bar{P}_s \leq \alpha \prod_{i=1}^{M} \frac{1}{1+\beta\bar{\gamma}_i/2}. \tag{7.33}$$

In the limit of high SNR and assuming that the γ_i are identically distributed with $\bar{\gamma}_i = \bar{\gamma}$, we have

$$\bar{P}_s \approx \alpha \left(\frac{\beta\bar{\gamma}}{2}\right)^{-M}. \tag{7.34}$$

Thus, at high SNR, the diversity order of MRC is M (the number of antennas) and so MRC achieves full diversity order.

MRC is the only diversity-combining technique that achieves full diversity order. However, it is shown in [8, Chap. 5.7] that performance of EGC is quite close to that of MRC, typically exhibiting less than 1 dB of power penalty. This is the price paid for the reduced complexity of using equal gains. A more extensive performance comparison between SC, MRC, and EGC can be found in [1, Chap. 9]. While MRC has the highest complexity of these three techniques, it is the most widely used in practice due to its superior performance.

Example 7.4: Compare the average probability of bit error with BPSK under MRC and EGC two-branch diversity, assuming i.i.d. Rayleigh fading with an average SNR of 10 dB on each branch.

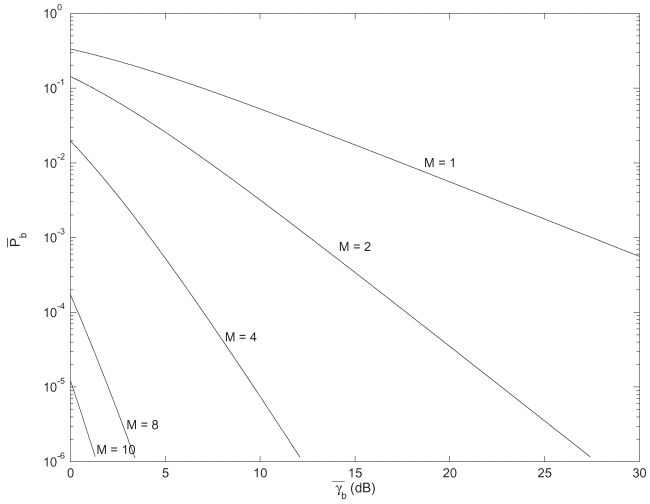


Figure 7.7: Average P_b for maximal-ratio combining with i.i.d. Rayleigh fading.

Solution: By (7.31), under MRC we have

$$\bar{P}_b = \left(\frac{1-\sqrt{10/11}}{2}\right)^2 (2+\sqrt{10/11}) = 1.60 \cdot 10^{-3}. \tag{7.35}$$

By (7.26), under EGC we have

$$\bar{P}_b = .5\left(1-\sqrt{1-\left(\frac{1}{11}\right)^2}\right) = 2.07 \cdot 10^{-3}. \tag{7.36}$$

Thus we see that the performance of MRC and EGC are almost the same.

7.2.6 Generalized Combining

The three most common diversity-combining schemes used in practice are MRC, EGC, and SC. MRC has the best performance and achieves full diversity order. However, MRC also has the highest complexity, as it requires coherent detection in each branch that determines the branch phase and amplitude. SC is the simplest of these three techniques as only one coherent detector is needed. The price paid for this simplicity is in performance, with diminishing diversity gain with each branch that is added. EGC has complexity and performance between the extremes of MRC and SC. Generalized Combining (GC), also referred to as Generalized Selection Combining or Hybrid Combining, utilizes SC to pick a subset of the branch signals which are then combined using MRC or EGC. In particular, a generalized combiner has $M > 2$ diversity branches of which it selects the $L > 1$ branches, $L < M$, with the highest SNR; these L branches are then combined using MRC or EGC. When MRC is used, the performance is better than an L-branch MRC system and worse than an M-branch MRC system, and similarly for EGC, due to the selection of the $L < M$ highest-SNR branches among the M available branches. Performance analysis of generalized combining under both MRC and EGC in different fading environments has been studied extensively (see [1, Chap. 9.11] and the references therein). This performance is best analyzed via the Moment Generating Function analysis for diversity systems, as will be explored in Section 7.4.

7.3 Transmitter Diversity

In transmit diversity there are multiple transmit antennas, and the transmit power is divided among these antennas. Transmit diversity is desirable in systems where more space, power, and processing capability are available on the transmit side than on the receive side, as best exemplified by the downlink in cellular systems. Transmit diversity design depends on whether or not the complex channel gain is known to the transmitter. When this gain is known, the system is quite similar to receiver diversity. However, without this channel knowledge, transmit diversity gain requires a combination of space and time diversity via a novel technique called the *Alamouti scheme* and its extensions. We now discuss transmit diversity under different assumptions about channel knowledge at the transmitter, assuming the channel gains are known at the receiver.

7.3.1 Channel Known at Transmitter

Consider a transmit diversity system with M transmit antennas and one receive antenna. We assume that the path gain $r_i e^{j\theta_i}$ associated with the ith antenna is known at the transmitter; this is referred to as having channel side information (CSI) at the transmitter, or CSIT. Let $s(t)$ denote the transmitted signal with average power P_t. The signal is multiplied by a complex gain $\alpha_i = a_i e^{-j\theta_i} (0 \leq a_i \leq 1)$ and then sent through the ith antenna. This complex multiplication performs both co-phasing and weighting relative to the channel gains. Because of the average power constraint P_t, the weights must satisfy $\sum_{i=1}^{M} a_i^2 = 1$. The weighted signals transmitted over all antennas are added "in the air", which leads to a received signal given by

$$r(t) = \sum_{i=1}^{M} a_i r_i s(t). \tag{7.37}$$

Let $N_0 B$ denote the noise power in the receiver and assume Nyquist pulses such that $BT_s = 1$.

Suppose we wish to set the branch weights to maximize received SNR. Using a similar analysis as in receiver MRC diversity, we see that the weights a_i that achieve the maximum SNR are given by

$$a_i = \frac{r_i}{\sqrt{\sum_{i=1}^{M} r_i^2}}, \tag{7.38}$$

and the resulting SNR is

$$\gamma_\Sigma = \frac{P_t}{N_0 B} \sum_{i=1}^{M} r_i^2 = \sum_{i=1}^{M} \gamma_i \tag{7.39}$$

for $\gamma_i = r_i^2 P_t / (N_0 B)$ equal to the branch SNR between the ith transmit antenna and the receive antenna. Thus we see that transmit diversity when the channel gains are known to the transmitter is very similar to receiver diversity with MRC: the received SNR is the sum of SNRs on each of the individual branches. In particular, if all antennas have the same gain $r_i = r$ then $\gamma_\Sigma = M r^2 P_t / (N_0 B)$, so there is an array gain of M corresponding to an M-fold increase in SNR over a single antenna transmitting with full power.

We again consider modulations whose error probability in AWGN can be approximated as $P_s = \alpha Q(\sqrt{\beta\gamma})$ for SNR γ, such as BPSK, MPSK, and MQAM. Using the Chernoff bound as in Section 7.2.5 we see that, for such modulations and static gains on each diversity branch,

$$P_s = \alpha Q(\sqrt{\beta\gamma_\Sigma}) \leq \alpha e^{-\beta\gamma_\Sigma/2} = \alpha e^{-\beta(\gamma_1 + \cdots + \gamma_M)/2}. \tag{7.40}$$

Integrating over the χ^2 distribution for γ_Σ yields the same bound as (7.33):

$$\bar{P}_s \leq \alpha \prod_{i=1}^{M} \frac{1}{1 + \beta\bar{\gamma}_i/2}. \tag{7.41}$$

In the limit of high SNR and assuming that the γ_i are identically distributed with $\bar{\gamma}_i = \bar{\gamma}$, we have

$$\bar{P}_s \approx \alpha \left(\frac{\beta \bar{\gamma}}{2} \right)^{-M}. \tag{7.42}$$

Thus, at high SNR, the diversity order of transmit diversity with MRC is M, so both transmit and receive MRC achieve full diversity order. The analysis for EGC and SC assuming transmitter channel knowledge is the same as under receiver diversity.

The complication of transmit diversity is to obtain the channel phase and, for SC and MRC, the channel gain, at the transmitter. These channel values can be measured at the receiver using a pilot technique and then fed back to the transmitter. Alternatively, in cellular systems with time division, the base station can measure the channel gain and phase on transmissions from the mobile to the base and then use these measurements when transmitting back to the mobile, since under time division the forward and reverse links are reciprocal.

7.3.2 Channel Unknown at Transmitter – The Alamouti Scheme

We now consider the same model as in Section 7.3.1 but assume here that the transmitter no longer knows the channel gains $r_i e^{j\theta_i}$, so there is no CSIT. In this case it is not obvious how to obtain diversity gain. Consider, for example, a naive strategy whereby for a two-antenna system we divide the transmit energy equally between the two antennas. Thus, the transmit signal on antenna i will be $s_i(t) = \sqrt{.5}s(t)$ for $s(t)$ the transmit signal with energy per symbol E_s. Assume that each antenna has a complex Gaussian channel gain $h_i = r_i e^{j\theta_i}$, $i = 1, 2$, with mean 0 and variance 1. The received signal is then

$$r(t) = \sqrt{.5}(h_1 + h_2)s(t). \tag{7.43}$$

Note that $h_1 + h_2$ is the sum of two complex Gaussian random variables and thus is itself a complex Gaussian, with mean equal to the sum of means (0) and variance equal to the sum of variances (2). Hence $\sqrt{.5}(h_1 + h_2)$ is a complex Gaussian random variable with mean 0 and variance 1, so the received signal has the same distribution as if we had just used one antenna with the full energy per symbol. In other words, we have obtained no performance advantage from the two antennas, since we could neither divide our energy intelligently between them nor obtain coherent combining via co-phasing.

Transmit diversity gain can be obtained even in the absence of channel information, given an appropriate scheme to exploit the antennas. A particularly simple and prevalent scheme for this diversity that combines both space and time diversity was developed by Alamouti in [11]. Alamouti's scheme is designed for a digital communication system with two-antenna transmit diversity. The scheme works over two symbol periods and it is assumed that the channel gain is constant over this time. Over the first symbol period, two different symbols s_1 and s_2 (each with energy $E_s/2$) are transmitted simultaneously from antennas 1 and 2, respectively. Over the next symbol period, symbol $-s_2^*$ is transmitted from antenna 1 and symbol s_1^* is transmitted from antenna 2, each again with symbol energy $E_s/2$.

Assume complex channel gains $h_i = r_i e^{j\theta_i}$, $i = 1, 2$, between the ith transmit antenna and the receive antenna. The received symbol over the first symbol period is $y_1 = h_1 s_1 + h_2 s_2 + n_1$ and the received symbol over the second symbol period is $y_2 = -h_1 s_2^* + h_2 s_1^* + n_2$, where n_i, $i = 1, 2$, is the AWGN sample at the receiver associated with the ith symbol transmission. We assume that the noise sample has mean 0 and power N.

The receiver uses these sequentially received symbols to form the vector $\mathbf{y} = [y_1 \;\; y_2^*]^T$ given by

$$\mathbf{y} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} = \mathbf{H}_A \mathbf{s} + \mathbf{n}, \tag{7.44}$$

where $\mathbf{s} = [s_1 \;\; s_2]^T, \mathbf{n} = [n_1 \;\; n_2^*]^T$, and

$$\mathbf{H}_A = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}. \tag{7.45}$$

Let us define the new vector $\mathbf{z} = \mathbf{H}_A^H \mathbf{y}$. The structure of $\mathbf{H}_A$ implies that

$$\mathbf{H}_A^H \mathbf{H}_A = (|h_1^2| + |h_2^2|)\mathbf{I}_2 \tag{7.46}$$

is diagonal and thus

$$\mathbf{z} = [z_1 \;\; z_2]^T = (|h_1^2| + |h_2^2|)\mathbf{I}_2 \mathbf{s} + \tilde{\mathbf{n}}, \tag{7.47}$$

where $\tilde{\mathbf{n}} = \mathbf{H}_A^H \mathbf{n}$ is a complex Gaussian noise vector with mean 0 and covariance matrix $E[\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H] = (|h_1^2| + |h_2^2|)N_0\mathbf{I}_2$. The diagonal nature of $\mathbf{z}$ effectively decouples the two symbol transmissions, so that each component of $\mathbf{z}$ corresponds to one of the transmitted symbols:

$$z_i = (|h_1^2| + |h_2^2|)s_i + \tilde{n}_i, \quad i = 1, 2. \tag{7.48}$$

The received SNR thus corresponds to the SNR for z_i given by

$$\gamma_i = \frac{(|h_1^2| + |h_2^2|)P_t}{2N_0 B}, \tag{7.49}$$

where the factor of 2 comes from the fact that s_i is transmitted using half the total power P_t. The received SNR is thus equal to the sum of SNRs on each branch divided by 2. By (7.48) the Alamouti scheme achieves a diversity order of 2 – the maximum possible for a two-antenna transmit system – despite the fact that channel knowledge is not available at the transmitter. However, by (7.49) it only achieves an array gain of 1, whereas MRC can achieve an array gain and a diversity gain of 2. The Alamouti scheme can be generalized for $M > 2$; this generalization falls into the category of orthogonal space-time block code design (see [12, Chap. 6.3.3] and [13, Chap. 7.4]).

7.4 Moment Generating Functions in Diversity Analysis

In this section we use the MGFs introduced in Section 6.3.3 to greatly simplify the analysis of average error probability under diversity. The use of MGFs in diversity analysis arises from the difficulty in computing the distribution $p_{\gamma_\Sigma}(\gamma)$ of the combiner SNR γ_Σ. Specifically, although the average probability of error and outage probability associated with diversity combining are given by the simple formulas (7.4) and (7.5), these formulas require integration over the distribution $p_{\gamma_\Sigma}(\gamma)$. This distribution is often not in closed form for an arbitrary number of diversity branches with different fading distributions on each branch, regardless of the combining technique that is used. In particular, the distribution for $p_{\gamma_\Sigma}(\gamma)$ is often in the form of an infinite-range integral, in which case the expressions for (7.4) and (7.5) become double integrals that can be difficult to evaluate numerically. Even when $p_{\gamma_\Sigma}(\gamma)$ is in closed form, the corresponding integrals (7.4) and (7.5) may not lead to closed-form solutions and may also be difficult to evaluate numerically. A large body of work over many decades has addressed approximations and numerical techniques for computing the integrals associated with average probability of symbol error for different modulations, fading distributions, and combining techniques (see [14] and the references therein). Expressing the average error probability in terms of the MGF for γ_Σ instead of its distribution often eliminates these integration difficulties. Specifically, when the diversity fading paths are independent but not necessarily identically distributed, the average error probability based on the MGF of γ_Σ is typically in closed form or consists of a single finite-range integral that can be easily computed numerically.

The simplest application of MGFs in diversity analysis is for coherent modulation with MRC, so this is treated first. We then discuss the use of MGFs in the analysis of average error probability under EGC, SC, and GC.

7.4.1 Diversity Analysis for MRC

The simplicity of using MGFs in the analysis of maximal-ratio combining stems from the fact that, as derived in Section 7.2.5, the combiner SNR γ_Σ is the sum of the γ_i, the branch SNRs:

$$\gamma_\Sigma = \sum_{i=1}^{M} \gamma_i. \tag{7.50}$$

As in the analysis of average error probability without diversity (Section 6.3.3), let us again assume that the probability of error in AWGN for the modulation of interest can be expressed either as an exponential function of γ_s, as in (6.78), or as a finite range integral of such a function, as in (6.79).

We first consider the case where $P_s = a\exp[-b\gamma]$ for SNR γ as given in (6.78), e.g., for DPSK with $a = .5$ and $b = 1$. Then the average probability of symbol error under MRC is

$$\bar{P}_s = \int_0^\infty a\exp[-b\gamma]p_{\gamma_\Sigma}(\gamma)\, d\gamma. \tag{7.51}$$

We assume that the branch SNRs are independent, so that their joint distribution becomes a product of the individual distributions: $p_{\gamma_1,\ldots,\gamma_M}(\gamma_1,\ldots,\gamma_M) = p_{\gamma_1}(\gamma_1)\ldots p_{\gamma_M}(\gamma_M)$. Using this factorization and substituting $\gamma = \gamma_1 + \cdots + \gamma_M$ in (7.51) yields

$$\bar{P}_s = a\underbrace{\int_0^\infty \int_0^\infty \cdots \int_0^\infty}_{M\text{-fold}} \exp[-b(\gamma_1 + \cdots + \gamma_M)]p_{\gamma_1}(\gamma_1)\ldots p_{\gamma_M}(\gamma_M)d\gamma_1\ldots d\gamma_M. \tag{7.52}$$

Now using the product forms $\exp[b(\gamma_1 + \cdots + \gamma_M)] = \prod_{i=1}^M \exp[b\gamma_i]$ and $p_{\gamma_1}(\gamma_1)\ldots p_{\gamma_M}(\gamma_M) = \prod_{i=1}^M p_{\gamma_i}(\gamma_i)$ in (7.52) yields

$$\bar{P}_s = a\underbrace{\int_0^\infty \int_0^\infty \cdots \int_0^\infty}_{M\text{-fold}} \prod_{i=1}^{M} \exp[-b\gamma_i]p_{\gamma_i}(\gamma_i)d\gamma_i. \tag{7.53}$$

Finally, switching the order of integration and multiplication in (7.53) yields our desired final form:

$$\bar{P}_s = a\prod_{i=1}^{M}\int_0^\infty \exp[-b\gamma_i]p_{\gamma_i}(\gamma_i)d\gamma_i = a\prod_{i=1}^{M}\mathcal{M}_{\gamma_i}(-b), \tag{7.54}$$

where $\mathcal{M}_{\gamma_i}(s)$ is the MGF of the fading distribution for the ith diversity branch as given by (6.74), (6.75), and (6.76) for (respectively) Rayleigh, Rician, and Nakagami fading. Thus, the average probability of symbol error is just the product of MGFs associated with the SNR on each branch.

We now consider modulations where $P_s = \alpha Q(\sqrt{\beta\gamma})$ in AWGN for γ the SNR. Using the alternate Q-function representation we get from (6.84) that

$$\begin{aligned}
\bar{P}_s &= \int_0^\infty \int_0^{\pi/2} a\exp[-b(\phi)\gamma]d\phi\, p_{\gamma_\Sigma}(\gamma)d\gamma \\
&= \underbrace{\int_0^\infty \int_0^\infty \cdots \int_0^\infty}_{M\text{-fold}} \int_0^{\pi/2} a\prod_{i=1}^{M}\exp[-b(\phi)\gamma_i]p_{\gamma_i}(\gamma_i)\, d\gamma_i d\phi.
\end{aligned} \tag{7.55}$$

with $a = \alpha/\pi$ and $b(\phi) = .5\beta/\sin^2\phi$; Again, switching the order of integration and multiplication yields our desired final form:

$$\bar{P}_s = \frac{\alpha}{\pi}\int_0^{\pi/2}\prod_{i=1}^{M}\int_0^{\infty}\exp\left[\frac{-.5\beta}{\sin^2\phi}\gamma_i\right]p_{\gamma_i}(\gamma_i)d\gamma_i dx = \frac{\alpha}{\pi}\int_0^{\pi/2}\prod_{i=1}^{M}\mathcal{M}_{\gamma_i}\left(\frac{-.5\beta}{\sin^2\phi}\right)\,d\phi. \tag{7.56}$$

Thus, the average probability of symbol error is just a single finite-range integral of the product of MGFs associated with the SNR on each branch. The simplicity of (7.54) and (7.56) are quite remarkable, given that these expressions apply for any number of diversity branches and any type of fading distribution on each branch (as long as the branch SNRs are independent).

We now apply these general results to specific modulations and fading distributions. Let us first consider DPSK, where $P_b(\gamma_b) = .5e^{-\gamma_b}$ in AWGN is in the form of (6.78) with $a = 1/2$ and $b = 1$. Then, by (7.54), the average probability of bit error in DPSK under M-fold MRC diversity is

$$\bar{P}_b = \frac{1}{2}\prod_{i=1}^{M}\mathcal{M}_{\gamma_i}(-1). \tag{7.57}$$

Note that this reduces to the probability of average bit error without diversity given by (6.71) for $M = 1$.

Example 7.5: Compute the average probability of bit error for DPSK modulation under three-branch MRC, assuming i.i.d. Rayleigh fading in each branch with $\bar{\gamma}_1 = 15$ dB and $\bar{\gamma}_2 = \bar{\gamma}_3 = 5$ dB. Compare with the case of no diversity for $\bar{\gamma} = 15$ dB.

Solution: From (6.74), $\mathcal{M}_{\gamma_i}(s) = (1 - s\bar{\gamma}_i)^{-1}$. Using this MGF in (7.57) with $s = -1$ yields

$$\bar{P}_b = \frac{1}{2}\frac{1}{1+10^{1.5}}\left(\frac{1}{1+10^{.5}}\right)^2 = 8.85\cdot 10^{-4}. \tag{7.58}$$

With no diversity we have

$$\bar{P}_b = \frac{1}{2(1+10^{1.5})} = 1.53\cdot 10^{-2}. \tag{7.59}$$

This indicates that additional diversity branches can significantly reduce average BER, even when the SNR on these branches is somewhat low.

Example 7.6: Compute the average probability of bit error for DPSK modulation under three-branch MRC, assuming: Nakagami fading in the first branch with $m = 2$ and $\bar{\gamma}_1 = 15$ dB; Rician fading in the second branch with $K = 3$ and $\bar{\gamma}_2 = 5$ dB; and Nakagami fading in the third branch with $m = 4$ and $\bar{\gamma}_3 = 5$ dB. Compare with the results of the prior example.

Solution: From (6.75) and (6.76), for Nakagami fading $\mathcal{M}_{\gamma_i}(s) = (1 - s\bar{\gamma}_i/m)^{-m}$ and for Rician fading

$$\mathcal{M}_{\gamma_s}(s) = \frac{1+K}{1+K-s\bar{\gamma}_s}\exp\left[\frac{Ks\bar{\gamma}_s}{1+K-s\bar{\gamma}_s}\right]. \tag{7.60}$$

Using these MGFs in (7.57) with $s = -1$ yields

$$\bar{P}_b = \frac{1}{2}\left(\frac{1}{1+10^{1.5}/2}\right)^2\frac{4}{4+10^{.5}}\exp\left[\frac{-3\cdot 10^{.5}}{4+10^{.5}}\right]\left(\frac{1}{1+10^{.5}/4}\right)^4 = 2.56\cdot 10^{-5}, \tag{7.61}$$

which is more than an order of magnitude lower than the average error probability under i.i.d. Rayleigh fading with the same branch SNRs derived in the previous problem. This indicates that Nakagami and Rician fading are much more benign distributions than Rayleigh, especially when multiple branches are combined under MRC. This example also illustrates the power of the MGF approach: computing average probability of error when the branch SNRs follow different distributions simply consists of multiplying together different functions in closed form; the result is then also in closed form. In contrast, computing the distribution of the sum of random variables from different families involves the convolution of their distributions, which rarely leads to a closed-form distribution.

For BPSK we see from (6.54) that P_b has the same form as (6.79) with $a = 1/\pi$, and $b(\phi) = 1/\sin^2\phi$. Thus we obtain the average bit error probability for BPSK with M-fold diversity as

$$\bar{P}_b = \frac{1}{\pi}\int_0^{\pi/2}\prod_{l=1}^{M}\mathcal{M}_{\gamma_i}\left(-\frac{1}{\sin^2\phi}\right)d\phi. \tag{7.62}$$

Similarly, if $P_s = \alpha Q(\sqrt{\beta\gamma_s})$ then the average probability of error in fading with M-branch diversity is given in (7.56) as

$$\bar{P}_s = \frac{\alpha}{\pi}\int_0^{\pi/2}\prod_{i=1}^{M}\mathcal{M}_{\gamma_i}\left(-\frac{.5\beta}{\sin^2\phi}\right)d\phi. \tag{7.63}$$

If the branch SNRs are i.i.d. then this expression simplifies to

$$\bar{P}_s = \frac{\alpha}{\pi}\int_0^{\pi/2}\left(\mathcal{M}_{\gamma}\left(-\frac{.5\beta}{\sin^2\phi}\right)\right)^{M}d\phi, \tag{7.64}$$

where $\mathcal{M}_\gamma(s)$ is the common MGF for the branch SNRs.

We can also use the MGF approach to determine the exact probability of symbol error for MPSK in fading. For this analysis, i.e. in equations (7.67)-(7.68), we use N to denote the number of branches and M for the signal constellation size. This analysis extends the derivation of (6.86) to multiple branches. As in that derivation, we note that Craig's formula for P_s of MPSK in AWGN is in the desired form (6.79) for MGF analysis. Using this form, we get the average symbol error probability in N-branch diversity as

$$\begin{aligned}
\bar{P}_s &= \int_0^{\infty}\int_0^{(M-1)\pi/M} a\exp[-b(\phi)\gamma]d\phi\, p_{\gamma_\Sigma}(\gamma)d\gamma & (7.65)\\
&= \underbrace{\int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty}}_{N\text{-fold}}\int_0^{(M-1)\pi/M} a\prod_{i=1}^{N}\exp[-b(\phi)\gamma_i]p_{\gamma_i}(\gamma_i)\,d\gamma_i d\phi & (7.66)\\
&= \frac{1}{\pi}\int_0^{(M-1)\pi/M}\prod_{i=1}^{N}\mathcal{M}_{\gamma_i}\left(-\frac{\beta_M}{\sin^2\phi}\right)d\phi, & (7.67)
\end{aligned}$$

for $\beta_M = \sin^2(\pi/M)$. For i.i.d. fading this simplifies to

$$\bar{P}_s = \frac{1}{\pi}\int_0^{(M-1)\pi/M}\left(\mathcal{M}_{\gamma}\left(-\frac{\beta_M}{\sin^2\phi}\right)\right)^{N}d\phi. \tag{7.68}$$

Example 7.7: Find an expression for the average symbol error probability for 8-PSK modulation for two-branch

MRC combining, where each branch is Rayleigh fading with average SNR of 20 dB.

Solution: The MGF for Rayleigh fading is $\mathcal{M}_{\gamma_i}(s) = (1 - s\bar{\gamma}_i)^{-1}$. Using this MGF in (7.68) with $s = -(\sin^2 \pi/8)/\sin^2 \phi$ and $\bar{\gamma} = 100$ yields

$$\bar{P}_s = \frac{1}{\pi} \int_0^{7\pi/8} \left(\frac{1}{1 + (100 \sin^2 \pi/8)/\sin^2 \phi} \right)^2 d\phi. \tag{7.69}$$

This expression does not lead to a closed-form solution and so must be evaluated numerically, which results in $\bar{P}_s = 1.56 \cdot 10^{-3}$.

We can use similar techniques to extend the derivation of the exact error probability for MQAM in fading, given by (6.93), to include MRC. As with MPSK, we let N denote the number of branches and M the constellation size in the expression (7.70) for $\bar{P}_s$. For this analysis we first integrate the expression for P_s in AWGN, expressed in (6.92) using the alternate representation of Q and Q^2, over the distribution of γ_Σ. Since $\gamma_\Sigma = \sum_i \gamma_i$ and the SNRs are independent, the exponential function and distribution in the resulting expression can be written in product form. Then we use the same reordering of integration and multiplication as in the MPSK derivation. The resulting average probability of symbol error for MQAM modulation with MRC combining of N branches is given by

$$\begin{aligned} \bar{P}_s =& \frac{4}{\pi} \left(1 - \frac{1}{\sqrt{M}} \right) \int_0^{\pi/2} \prod_{i=1}^{N} \mathcal{M}_{\gamma_i} \left(-\frac{.5\beta}{\sin^2 \phi} \right) d\phi \\ & - \frac{4}{\pi} \left(1 - \frac{1}{\sqrt{M}} \right)^2 \int_0^{\pi/4} \prod_{i=1}^{N} \mathcal{M}_{\gamma_i} \left(-\frac{.5\beta}{\sin^2 \phi} \right) d\phi. \end{aligned} \tag{7.70}$$

More details on the use of MGFs to obtain average probability of error under M-fold MRC diversity for a broad class of modulations can be found in [1, Chap. 9.2].

7.4.2 Diversity Analysis for EGC, SC, and GC

Moment generating functions are less useful in the analysis of EGC and SC than in MRC. The reason is that, with MRC, $\gamma_\Sigma = \sum_i \gamma_i$ and so $\exp[-b\gamma_\Sigma] = \prod_i \exp[-b\gamma_i]$. This factorization leads directly to the simple formulas whereby probability of symbol error is based on a product of MGFs associated with each of the branch SNRs. Unfortunately, neither EGC nor SC leads to this type of factorization. However, working with the MGF of γ_Σ can sometimes lead to simpler results than working directly with its distribution. This is illustrated in [1, Chap. 9.3.3]: the exact probability of symbol error for MPSK with EGC is obtained based on the characteristic function associated with each branch SNR, where the characteristic function is just the MGF evaluated at $s = j2\pi f$ (i.e., it is the Fourier transform of the distribution). The resulting average error probability, given by [1, eq. (9.78)], is a finite-range integral over a sum of closed-form expressions and thus is easily evaluated numerically.

In the case of GC where the $L < M$ fading branches with the L largest SNRs are combined using MRC, the MGF approach combined with order statistics can be used for performance analysis. Let γ_i, $i = 1, \ldots, M$, denote the SNR on the ith branch. We can order these random variables from smallest to largest as $\gamma_{(1)}, \ldots, \gamma_{(M)}$, $\gamma_{(i)} \leq \gamma_{(i+1)}$, with $\gamma_{(1)} = \min[\gamma_1, \gamma_2, \ldots, \gamma_M]$ and $\gamma_{(M)} = \max[\gamma_1, \gamma_2, \ldots, \gamma_M]$. Note that these ordered random variables are not i.i.d. The largest L of these values are $\gamma_{(M)}, \gamma_{(M-1)}, \ldots, \gamma_{(M-L+1)}$ where, by definition, $\gamma_{(M)} \geq \gamma_{(M-1)} \geq, \ldots, \geq \gamma_{(M-L+1)}$. For GC using MRC to combine diversity branches with these L largest SNRs, its SNR is

$$\gamma_\Sigma = \sum_{i=M-L+1}^{M} \gamma_{(i)}. \tag{7.71}$$

The distribution of γ_Σ is difficult to obtain directly for general distributions on the branch SNRs, as is the case for MRC of all branches (i.e. when $M = L$). However, as with MRC, we can obtain the distribution of γ_Σ through the MGF approach as follows.

Since γ_i, $i = 1, \ldots, M$ are i.i.d., the joint distribution of $\gamma_{(M)}, \ldots, \gamma_{(M-L+1)}$, i.e.,

$$p_{\gamma_{(M)},\ldots,\gamma_{(M-L+1)}}(x_{(M)}, \ldots, x_{(M-L+1)}) = p(\gamma_{(M)} = x_{(M)}, \ldots, \gamma_{(M-L+1)} = x_{(M-L+1)}) \tag{7.72}$$

can be shown to equal [1, eq. (9.419)]

$$p_{\gamma_{(M)},\ldots,\gamma_{(M-L+1)}}(x_{(M)}, \ldots, x_{(M-L+1)}) = \frac{M!}{(M-L)!} \left[P_\gamma(x_{(M-L+1)})\right]^{M-L} \prod_{i=1}^{L} p_\gamma(x_{(M-L+i)}), \tag{7.73}$$

where $p_\gamma(x) = p(\gamma = x)$ is the pdf of the branch SNR and $P_\gamma(x) = p(\gamma < x)$ is the cdf of the branch SNR. For example, in i.i.d. Rayleigh fading, $p_\gamma(\gamma)$ is exponential with mean $\overline{\gamma}$ and $P_\gamma(\gamma) = 1 - e^{-\gamma/\overline{\gamma}}$.

The MGF of γ_Σ given by (7.71) is

$$M_{\gamma_\Sigma}(s) = E_{\gamma_\Sigma}\left[e^{s\gamma_\Sigma}\right], \tag{7.74}$$

where $E_{\gamma_\Sigma}[\cdot]$ is the expectation with respect to the distribution of γ_Σ as given by (7.73). A closed-form expression for $M_{\gamma_\Sigma}(s)$ can then be obtained by writing γ_Σ in terms of the $L-1$ differences or *spaces* x_i between the L largest branch SNRs, given by $x_i = \gamma_{(M-i)} - \gamma_{(M-i-1)}, i = 0, \ldots, L-1$. For i.i.d. Rayleigh fading on each of the M total branches with branch SNR $\overline{\gamma}$ these spaces have been shown in [1, Theorem 9.1] to follow the exponential distribution

$$p_{x_i}(x) = \frac{1}{\overline{\gamma}} e^{-ix/\overline{\gamma}}. \tag{7.75}$$

Expressing the MGF (7.71) in terms of these x_i's yields, after some manipulation, a closed-form result for the MGF of GC with MRC of the L highest-SNR branches given by

$$M_{\gamma_\Sigma}(s) = (1 - s\overline{\gamma})^{-L+1} \prod_{i=L}^{M} \left(1 - \frac{s\overline{\gamma}L}{i}\right)^{-1}. \tag{7.76}$$

Differentiating this expression yields the pdf of γ_Σ which can then be used to obtain P_{out} and $\bar{P}_s$. More details can be found in [1, Chap. 9.11.2.1].

7.4.3 Diversity Analysis for Noncoherent and Differentially Coherent Modulation

A similar MGF approach to determining the average symbol error probability of noncoherent and differentially coherent modulations with diversity combining is presented in [1, 15]. This approach differs from that of the coherent modulation case in that it relies on an alternate form of the Marcum Q-function instead of the Gaussian Q-function, since the BER of noncoherent and differentially coherent modulations in AWGN are given in terms of the Marcum Q-function. Otherwise the approach is essentially the same as in the coherent case, and it leads to BER expressions involving a single finite-range integral that can be readily evaluated numerically. More details on this approach are found in [1] and [15].

Chapter 7 Problems

1. Find the outage probability of QPSK modulation at $P_s = 10^{-3}$ for a Rayleigh fading channel with SC diversity for $M = 1$ (no diversity), $M = 2$, and $M = 3$. Assume branch SNRs of $\bar{\gamma}_1 = 10\,\text{dB}$, $\bar{\gamma}_2 = 15\,\text{dB}$, and $\bar{\gamma}_3 = 20\,\text{dB}$.

2. This problem explores diversity gain under both outage probability and average probability of error metrics, and shows they are the same for extreme SNR values. Consider an M-branch diversity system with independent Rayleigh fading on each path. The average SNR at each branch is given by $\overline{\gamma}$. Assume DPSK transmission.

   (a) Defining outage probability for $M > 1$ as the probability of the instantaneous SNR γ at all the receiver antennas being less than a constant c, compute $-\frac{\log(\texttt{outage probability})}{\log(\overline{\gamma})}$ as $\overline{\gamma} \to \infty$. Does it depend on c ?

   (b) Now consider an alternative definition of diversity gain based on average error probability as

   $$-\frac{\log(\texttt{average probability of error with MRC})}{\log(\overline{\gamma})}.$$

   For $M > 1$, evaluate this as $\overline{\gamma} \to \infty$.

   (c) Now evaluate the diversity gain according to both the definitions as $\overline{\gamma} \to 0$. Are these definitions of diversity gain more suitable at high or low average SNR to describe the gains from having $M > 1$ independent fading paths?

3. This problem explores the derivation of (7.12). Show that the average SNR of the combiner output under SC diversity in independent and identically distributed Rayleigh fading is

   $$\bar{\gamma}_\Sigma = \int_0^\infty \frac{\gamma M}{\bar{\gamma}} [1 - e^{-\gamma/\bar{\gamma}}]^{M-1} e^{-\gamma/\bar{\gamma}} d\gamma = \bar{\gamma} \sum_{i=1}^{M} \frac{1}{i}$$

   by expressing $[1 - e^{-\gamma/\bar{\gamma}}]^{M-1}$ using its binomial expansion, performing the integration in the resulting expression, and applying induction to this simplified expression to get the final result.

4. This problem uses order statistics as an alternate method to derive the result of (7.12). For a set of M identically and independently distributed random variables, $X_1, X_2, \ldots, X_M$, the order statistics of these variables are given by $X_{(1)}, X_{(2)}, \ldots, X_{(M)}$ where

   $$X_{(1)} = \min[X_1, X_2, \ldots, X_M] \qquad X_{(M)} = \max[X_1, X_2, \ldots, X_M].$$

   In general, $X_{(i)}$ denotes the ith largest value of $X_1, X_2, \ldots, X_n$, with $X_{(1)} \leq X_{(2))} \leq \ldots \leq X_{(M)}$. A well-known result in order statistics states that if $X_1, \ldots X_M$ are i.i.d random variables drawn from an exponential distribution with mean $\overline{X}$, then the order statistics of this collection of variables are given by

   $$X_{(i)} = \overline{X} \left(\sum_{j=1}^{i} \frac{Z_j}{n - j + 1} \right), \tag{7.77}$$

   where $Z_1, \ldots, Z_M$ are independent exponentially distributed random variables with mean 1.

   (a) For selection-combining (SC) of M i.i.d. branches under Rayleigh fading where each branch has the same average received power $\bar{\gamma}$, use the result of (7.77) to derive (7.12).

(b) Suppose that $\bar{\gamma} = 20$ dB. Use order statistics of i.i.d. exponentially-distributed random variables to find P_{out} for the cases listed below assuming BPSK modulation with $P_b = 10^{-4}$.

i. $M = 1$ (no diversity)

ii. $M = 3$. Also provide an expression for P_{out} in terms of the cdf of $X_{(3)}$.

(c) Suppose that instead of an SC diversity scheme, we have a single transmitter sending to three users, each with a single antenna. Assume BPSK modulation is sent over each user's single-antenna channel and these channels experience i.i.d. Rayleigh fading with SNR $\bar{\gamma} = 20$ dB. A user is in outage if $P_b > 10^{-4}$. Compute the probability associated with each of the following events and express these probabilities in terms of the cdfs of the variables $X_{(1)}, X_{(2)}$, and $X_{(3)}$:

i. No users are in outage.

ii. One or fewer users are in outage.

iii. Two or fewer users are in outage.

iv. All users are in outage.

5. Plot the distribution $p_{\gamma_\Sigma}(\gamma)$ given by (7.11) for the selection combiner SNR in Rayleigh fading with M-branch diversity assuming $M = 1, 2, 4, 8$, and 10. Assume that each branch has an average SNR of 10 dB. Your plot should be linear on both axes and should focus on the range of linear γ values $0 \leq \gamma \leq 60$. Discuss how the distribution changes with increasing M and why this leads to lower probability of error.

6. Derive the average probability of bit error for DPSK under SC with i.i.d. Rayleigh fading on each branch as given by (7.13).

7. Derive a general expression for the cdf of the two-branch SSC output SNR for branch statistics that are not i.i.d., and show that it reduces to (7.14) for i.i.d. branch statistics. Evaluate your expression assuming Rayleigh fading in each branch with different average SNRs $\bar{\gamma}_1$ and $\bar{\gamma}_2$.

8. Derive the average probability of bit error for DPSK under SSC with i.i.d. Rayleigh fading on each branch as given by (7.21).

9. Plot the average probability of bit error for DPSK under MRC with $M = 2, 3$, and 4, assuming i.i.d. Rayleigh fading on each branch and an average branch SNR ranging from 0 dB to 20 dB.

10. Show that the weights a_i maximizing γ_Σ under receiver diversity with MRC are $a_i^2 = r_i^2/N_0$ for $N_0/2$ the common noise PSD on each branch. Also show that, with these weights, $\gamma_\Sigma = \sum_i \gamma_i$.

11. This problem illustrates that, because of array gain, you can get performance gains from diversity combining even without fading. Consider an AWGN channel with N-branch diversity combining and $\gamma_i = 10\,\text{dB}$ per branch. Assume MQAM modulation with $M = 4$ and use the approximation $P_b = .2e^{-1.5\gamma/(M-1)}$ for bit error probability, where γ is the received SNR.

(a) Find P_b for $N = 1$.

(b) Find N so that, under MRC, $P_b < 10^{-6}$.

12. Derive the average probability of bit error for BPSK under MRC with i.i.d. Rayleigh fading on each branch as given by (7.31).

13. Derive the average probability of bit error for BPSK under EGC with i.i.d. Rayleigh fading on each branch as given by (7.26).

14. Compare the average probability of bit error for BPSK modulation under no diversity, two-branch SC, two-branch SSC with $\gamma_T = \gamma_0$, two-branch EGC, and two-branch MRC. Assume i.i.d. Rayleigh fading on each branch with equal branch SNRs of 10 dB and of 20 dB. How does the relative performance change as the branch SNRs increase?

15. Plot the average probability of bit error for BPSK under both MRC and EGC assuming two-branch diversity with i.i.d. Rayleigh fading on each branch and average branch SNR ranging from 0 dB to 20 dB. What is the maximum dB penalty of EGC as compared to MRC?

16. Compare the outage probability of BPSK modulation at $P_b = 10^{-3}$ under MRC versus EGC, assuming two-branch diversity with i.i.d. Rayleigh fading on each branch and average branch SNR $\bar{\gamma} = 10\,\mathrm{dB}$.

17. Compare the average probability of bit error for BPSK under MRC versus EGC, assuming two-branch diversity with i.i.d. Rayleigh fading on each branch and average branch SNR $\bar{\gamma} = 10\,\mathrm{dB}$.

18. Consider a GC diversity system with M receiver branches and L receivers, whereby the combiner chooses the L highest SNR signals from the M branches and then combines them using MRC. Assume the SNR on the ith branch is γ_i which is i.i.d. distributed uniformly between 0 and 1 in linear units. Denote the SNR for the combiner output as γ_Σ; the outage probability for threshold γ_0 is $P_{out} = p(\gamma_\Sigma < \gamma_0)$. Assume that there are $M = 3$ branches and the threshold is $\gamma_0 = 1$.

    (a) Suppose $L = 1$, state the form of diversity combining for this special case and compute P_{out}.

    (b) Repeat part (a) if $L = 3$.

    (c) Suppose $L = 2$. Show that $p(\gamma_{\max} < \gamma_0/2) \leq P_{out} = p(\gamma_\Sigma < \gamma_0) \leq p(\gamma_{\min} < \gamma_0/2)$ for $\gamma_{\min} = \min_i \gamma_i, \gamma_{\max} = \max_i \gamma_i$. Compute upper and lower bounds for P_{out}.

19. Compute the average BER of a channel with two-branch transmit diversity under the Alamouti scheme, assuming Rayleight fading in each branch with an average branch SNR of 10 dB.

20. Consider a fading distribution $p(\gamma)$ where $\int_0^\infty p(\gamma)e^{-x\gamma}d\gamma = (1 + 2\bar{\gamma}x)^{-1/2}$. Find the average P_b for a BPSK modulated signal where (a) the receiver has two-branch diversity with MRC combining and (b) each branch has an average SNR of 10 dB and experiences independent fading with distribution $p(\gamma)$.

21. Consider a fading channel that is BPSK modulated and has three-branch diversity with MRC, where each branch experiences independent fading with an average received SNR of 15 dB. Compute the average BER of this channel for Rayleigh fading and for Nakagami fading with $m = 2$. *Hint*: Using the alternate Q-function representation greatly simplifies this computation, at least for Nakagami fading.

22. Plot the average probability of error as a function of branch SNR for a two-branch MRC system with BPSK modulation, where the first branch has Rayleigh fading and the second branch has Nakagami-m fading with $m = 2$. Assume the two branches have the same average SNR; your plots should have this average branch SNR ranging from 5 dB to 20 dB.

23. Plot the average probability of error as a function of branch SNR for an M-branch MRC system with 8-PSK modulation for $M = 1, 2, 4, 8$. Assume that each branch has Rayleigh fading with the same average SNR. Your plots should have this SNR ranging from 5 dB to 20 dB.

24. Derive the average probability of symbol error for MQAM modulation under MRC diversity given by (7.70) from the probability of error in AWGN (6.92) by utilizing the alternate representation of Q and Q^2.

25. Compare the average probability of symbol error for 16-PSK and 16-QAM modulation, assuming three-branch MRC diversity with Rayleigh fading on the first branch and Rician fading on the second and third branches with $K = 2$. Assume equal average branch SNRs of 10 dB.

26. Plot the average probability of error as a function of branch SNR for an M-branch MRC system with 16-QAM modulation for $M = 1, 2, 4, 8$. Assume that each branch has Rayleigh fading with the same average SNR. Your plots should have an SNR ranging from 5 dB to 20 dB.

27. MRC:

    (a) Show that the weights a_i maximizing γ_Σ under receiver diversity with MRC are $a_i^2 = r_i^2/N_0B$ for $N_0B/2$ the common noise PSD on each branch. Also show that, with these weights, $\gamma_\Sigma = \sum_i \gamma_i$, as stated in (7.28) of the reader.

    (b) Consider a fading distribution $p(\gamma)$ where $\int_0^\infty p(\gamma)e^{-x\gamma}\,\mathrm{d}\gamma = (1 + 2\bar{\gamma}x)^{-1/2} \quad \forall x > -1/2\bar{\gamma}$. Find the average P_b for a BPSK modulated signal where the receiver has 2-branch diversity with MRC combining, and each branch has an average SNR of 10 dB and experiences independent fading with distribution $p(\gamma)$.

Bibliography

[1] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, Wiley, New York, 2nd Ed., 2004.

[2] A. F. Molisch, *Wireless Communications: From Fundamentals to Beyond 5G*, Wiley-IEEE Press, 3rd ed. 2022.

[3] J. Winters, "Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading," *IEEE Trans. Veh. Tech.*, pp. 1740–51, November 1993.

[4] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[5] M. Blanco and K. Zdunek, "Performance and optimization of switched diversity systems for the detection of signals with Rayleigh fading," *IEEE Trans. Commun.*, pp. 1887–95, December 1979.

[6] A. Abu-Dayya and N. Beaulieu, "Switched diversity on microcellular Ricean channels," *IEEE Trans. Veh. Tech.*, pp. 970–6, November 1994.

[7] A. Abu-Dayya and N. Beaulieu, "Analysis of switched diversity systems on generalized-fading channels," *IEEE Trans. Commun.*, pp. 2959–66, November 1994.

[8] M. Yacoub, *Principles of Mobile Radio Engineering*, CRC Press, Boca Raton, FL, 1993.

[9] T. Eng, N. Kong and L. B. Milstein, "Comparison of diversity combining techniques for Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 44, no. 9, pp. 1117-1129, Sept. 1996.

[10] M.-S. Alouini and M. K. Simon, "An MGF-based performance analysis of generalized selection combining over Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 401-415, March 2000.

[11] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, pp. 1451–8, October 1998.

[12] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

[13] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.

[14] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communications over generalized fading channels," *Proc. IEEE*, pp. 1860–77, September 1998.

[15] M. K. Simon and M.-S. Alouini, “A unified approach for the probability of error for noncoherent and differentially coherent modulations over generalized fading channels,” *IEEE Trans. Commun.*, pp. 1625–38, December 1998.

Chapter 8

Coding for Wireless Channels

Coding allows bit errors introduced by transmission of a modulated signal through a wireless channel to be either detected or corrected by a decoder in the receiver. Coding can be considered as the embedding of signal constellation points in a higher-dimensional signaling space than is needed for communications. By going to a higher-dimensional space, the distance between points can be increased, which provides for better error correction and detection. As shown in Figure 8.1, encoding is typically performed on the information bits to be transmitted over the channel prior to modulation. Decoding is performed on the (possibly corrupted) coded bits after demodulated. The effects of the modulation, transmission, and demodulation on the coded bits can be modeled as a binary channel.

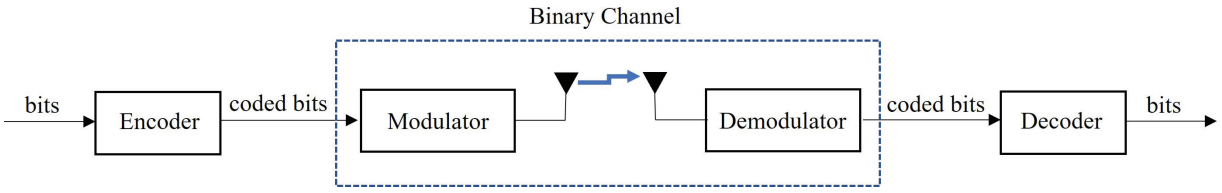


Figure 8.1: Coding Block Diagram.

Codes designed for AWGN channels typically do not work well on fading channels because they cannot correct for long error bursts that occur in deep fading. Codes for fading channels are mainly based on using an AWGN channel code combined with interleaving, but the criterion for the code design changes to provide fading diversity. Other coding techniques to combat performance degradation due to fading include unequal error protection codes and joint source and channel coding. All of these techniques will be described in this chapter.

We first provide an overview of code design for AWGN channels, along with basic code design parameters such as code rate, spectral efficiency, and minimum distance. This section will also discuss coding gain in AWGN and its theoretical maximum, the Shannon limit. Sections 8.2 and 8.4 describe algebraic block and convolutional code designs. These codes, along with their concatenation as described in Section 8.5, dominated the field of channel coding in the first few decades after Shannon's landmark 1948 paper on channel capacity and the random codes that achieve it. Yet approaching the Shannon limit with such codes proved elusive. The coding breakthrough needed to approach the Shannon limit required a new paradigm, that of "codes on sparse graphs", of which the widely-used turbo and low-density parity-check (LDPC) codes are special cases. These codes along with their soft-decision iterative decoding techniques are described in Sections 8.6 and 8.7, respectively. Turbo and LDPC codes exhibit near-capacity performance with reasonable decoding complexity at moderate-to-high data rates. For low data rates applications, polar codes - described in Section 8.8 - have emerged as a compelling error-correction

code due to their excellent performance with low-complexity decoding. Polar codes are a type of block code whose performance approaches the Shannon limit at large blocklengths, but practical designs generally approach this limit only at low data rates.

Wireless channels are generally restricted to a given bandwidth and suffer from fading as well as AWGN. Coded modulation was invented in the late 1970s as a technique to obtain error correction in bandlimited channels through a joint design of the modulation and coding. We will discuss the basic design principles behind trellis and more general lattice coded modulation along with their performance in Section 8.10. Code designs for fading channels are covered in Section 8.11. These techniques combine block or convolutional codes designed for AWGN channels with interleaving, with the AWGN code design metric modified to incorporate maximum fading diversity. Diversity gains can also be obtained by combining coded modulation with symbol or bit interleaving, with bit interleaving providing much higher diversity gain. Thus, coding combined with interleaving provides diversity gain in the same manner as other forms of diversity, with the diversity order built into the code design. The channel gain (or SNR) of a fading channel may change over the transmission of an encoded bit stream. Unequal error protection (UEP) and fountain codes, described in Section 8.12, are used in fading channels when the channel gain is variable or unknown at the transmitter. In UEP codes bits are prioritized, with high-priority bits encoded using a stronger error protection against deep fades than the low-priority bits. Thus, if the channel enters a deep fade during transmission of an encoded bit stream, the high priority bits can still be recovered. In fountain codes a block of bits is encoded into a long stream of encoded bits. The receiver decodes the stream of encoded bits until it has sufficiently many to decode the transmitted bits; the number of encoded bits needed depends on the channel gain.

Coding is a broad and deep subject that cannot be fully covered in a single book chapter. Indeed, there are many excellent books and tutorials providing in-depth treatments of the topics in this chapter [1, 2, **?**, 4]. The historical survey [5] maps out the evolution of coding from early algebraic and convolutional codes to near-capacity-achieving Turbo and LDPC codes, including their theoretical foundations, code design principles, and optimal decoding strategies.

8.1 Overview of Code Design

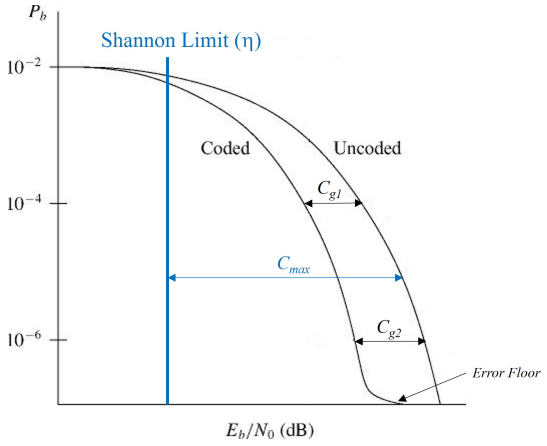


Figure 8.2: Coding gain in AWGN channels and the Shannon limit $(2^{\eta} - 1)/\eta$ for spectral efficiency η.

The main reason to apply error correction coding in a wireless system is to reduce the probability of bit or block error. The bit error probability P_b for a coded system is the probability that a bit is decoded in error. The block error probability P_{bl}, also called the packet error rate, is the probability that one or more bits in a block of

coded bits is decoded in error. Block error probability is useful for packet data systems where bits are encoded and transmitted in blocks. The amount of error reduction provided by a given code is typically characterized by its coding gain in AWGN and its diversity gain in fading.

Coding gain in AWGN is defined as the amount that the bit energy or signal-to-noise power ratio can be reduced under the coding technique for a given P_b or P_{bl}. We illustrate coding gain for P_b in Figure 8.2. We see in this figure that the gain C_{g1} at $P_b = 10^{-4}$ is less than the gain C_{g2} at $P_b = 10^{-6}$, and there is negligible coding gain at $P_b = 10^{-2}$. In fact, codes designed for high-SNR channels can have negative coding gain at low SNRs, since the extra redundancy of the code does not provide sufficient performance gain in P_b or P_{bl} at low SNRs to compensate for spreading the bit energy over multiple coded bits. The coding gain in AWGN is generally a function of the minimum Euclidean distance of the code, which equals the minimum distance in signal space between codewords or error events. Thus, codes designed for AWGN channels maximize their Euclidean distance for good performance.

In addition to coding gain, code performance in AWGN is commonly measured against the *Shannon limit*, which is obtained from the channel capacity formula $B\log_2(1 + \mathrm{SNR})$ bps. As discussed in Chapter 4.1, at data rates up to capacity, the capacity-achieving code has a probability of error that goes to zero, as indicated by the straight line in Figure 8.2. The Shannon limit thus indicates the best possible performance that any practical code can achieve over an AWGN channel. The Shannon limit is obtained as follows. Recall from Chapter 6.1.1 that $E_b/N_0 = \mathrm{SNR}/\eta$ for $\eta = R/B$ the spectral efficiency of the communication link in bps/Hz. Shannon capacity dictates that error probability approaches zero for any rate $R < B\log_2(1 + \mathrm{SNR})$ or, equivalently, that $\eta = \log_2(1 + \mathrm{SNR})$ is the maximum spectral efficiency a system can achieve for a given SNR with asymptotically small error probability. Inverting this expression yields a lower bound on the SNR required to achieve a given spectral efficiency with asymptotically small error probability given by $\mathrm{SNR} \geq 2^\eta - 1$. Using the relationship $E_b/N_0 = \mathrm{SNR}/\eta$, this gives a lower bound on the E_b/N_0 required to achieve a given spectral efficiency with asymptotically small error probability of

$$E_b/N_0 \geq \frac{2^\eta - 1}{\eta}. \tag{8.1}$$

This bound is often referred to as the *Shannon limit* for a given spectral efficiency η. In the limit of infinite bandwidth B or asymptotically small data rate R, $\eta = R/B$ approaches zero, and the Shannon limit is given by

$$\lim_{\eta \to 0} \frac{2^\eta - 1}{\eta} = -1.59 \text{ dB}. \tag{8.2}$$

The straight line in Figure 8.2 indicating the Shannon limit can be for a particular value of η to compare "apples to apples", i.e. to compare codes with spectral efficiency η against their best possible performance, in which case the Shannon limit is given by Equation 8.1. Alternatively, the Shannon limit can be given by the asymptotic limit of -1.59 dB given by Equation 8.2 if spectral efficiency of the code is a design parameter that can be reduced to approach this asymptotic limit. The maximum coding gain $C_{\max}$ that a code design can achieve is the difference in dB between the E_b/N_0 required to achieve a given error probability for uncoded transmission and the Shannon limit.

Error probability in AWGN with or without coding tends to fall off with SNR as a waterfall shape at low to moderate SNRs. This waterfall shape holds at all SNRs for uncoded systems and many coded systems. However some codes, in particular turbo and LDPC codes, exhibit error floors as SNR grows. The error floor, also shown in Figure 8.2, kicks in at a threshold SNR that depends on the code design. For SNRs above this threshold, the slope of the error probability curve decreases because minimum distance error events dominate code performance in this SNR regime.

For many codes, the error correction capability of a code does not come for free. This performance enhancement is paid for by increased complexity and – for block codes, convolutional codes, turbo codes, and LDPC codes

– by either a decrease in data rate or an increase in signal bandwidth. Consider a code with n coded bits for every k uncoded bits. This code effectively embeds a k-dimensional subspace into a larger n-dimensional space to provide larger distances between coded symbols. However, if the data rate through the channel is fixed at R_b, then the information rate for a code that uses n coded bits for every k uncoded bits is $(k/n)R_b$; that is, coding decreases the data rate by the fraction k/n. We can keep the information rate constant and introduce coding gain by decreasing the bit time by k/n. This typically results in an expanded bandwidth of the transmitted signal by n/k. Coded modulation uses a joint design of the code and modulation to obtain coding gain without this bandwidth expansion, as discussed in more detail in Section 8.9.

Codes designed for AWGN channels do not generally work well in fading owing to bursts of errors that cannot be corrected for. However, good performance in fading can be obtained by combining AWGN channel codes with interleaving and by designing the code to optimize its inherent diversity. The interleaver spreads out bursts of errors over time, so it provides a form of time diversity. This diversity is exploited by the inherent diversity in the code. In fact, codes designed in this manner exhibit performance similar to maximal-ratio combining (MRC) diversity. As discussed in Chapter 7.2.5, MRC with a diversity order of M reduces the average probability of error in fading from $c\bar{\gamma}^{-1}$ to $c\bar{\gamma}^{-M}$ for some constant c and average SNR $\bar{\gamma}$. As will be discussed in Section 8.10, the diversity order of coding with interleaving has a diversity order equal to the code's *minimum Hamming distance*, defined as the number of coded symbols that differ between different codewords or error events. Thus, coding combined with interleaving provides the same performance gain in fading as MRC with diversity order equal to the code's minimum Hamming distance.

8.2 Binary Linear Block Codes

Binary linear block codes are conceptually simple codes that are basically an extension of single-bit parity-check codes for error detection. A single-bit parity-check code is one of the most common forms of detecting transmission errors. This code uses one extra bit in a block of n data bits to indicate whether the number of l-bits in a block is odd or even. Thus, if a single error occurs, either the parity bit is corrupted or the number of detected 1-bits in the information bit sequence will be different from the number used to compute the parity bit; in either case the parity bit will not correspond to the number of detected 1-bits in the information bit sequence, so the single error is detected. Linear block codes extend this notion by using a larger number of parity bits to either detect more than one error or correct for one or more errors. Unfortunately, linear block codes – along with convolutional codes – trade their error detection or correction capability for either bandwidth expansion or a lower data rate, as we shall discuss in more detail.

In this section we focus on binary linear block codes, where the block code consists of coded bits taking a value of either 0 or 1. More specifically, a binary linear block code generates a block of n coded bits from k information bits. We call this an (n, k) binary block code. The coded bits are also called *codeword symbols*. The n codeword symbols can take on 2^n possible values corresponding to all possible combinations of the n binary bits. We select 2^k codewords from these 2^n possibilities to form the code, where each k bit information block is uniquely mapped to one of these 2^k codewords. The rate of the code is $R_c = k/n$ information bits per codeword symbol. If we assume that codeword symbols are transmitted across the channel at a rate of R_s symbols per second, then the information rate associated with an (n, k) block code is $R_b = R_c R_s = (k/n)R_s$ bits per second. Thus we see that block coding reduces the data rate compared to what we obtain with uncoded modulation by the code rate R_c.

A block code is called a *linear* code when the mapping of the k information bits to the n codeword symbols is a linear mapping. In order to describe this mapping and the corresponding encoding and decoding functions in more detail, we must first discuss properties of the vector space of binary n-tuples and its corresponding subspaces.

8.2.1 Binary Vector Spaces and Hamming Distance

The set of all binary n-tuples B_n is a vector space over the binary field, which consists of the two elements 0 and 1. All fields have two operations, addition and multiplication: for the binary field these operations correspond to binary addition (modulo 2 addition) and standard multiplication. A subset S of B_n is called a *subspace* if it satisfies the following conditions.

1. The all-zero vector is in S.
2. The set S is closed under addition; that is, if $S_i \in S$ and $S_j \in S$ then $S_i + S_j \in S$.

An (n, k) block code is linear if the 2^k length-n codewords of the code form a subspace of B_n. Thus, if $\mathbf{C}_i$ and $\mathbf{C}_j$ are two codewords in an (n, k) linear block code, then $\mathbf{C}_i + \mathbf{C}_j$ must form another codeword of the code.

Example 8.1: The vector space B_3 consists of all binary tuples of length 3:

$$B_3 = \{[000], [001], [010], [011], [100], [101], [110], [111]\}. \tag{8.3}$$

Note that B_3 is a subspace of itself, since it contains the all-zero vector and is closed under addition. Determine which of the following subsets of B_3 form a subspace:

- $A_1 = \{[000], [001], [100], [101]\}$;
- $A_2 = \{[000], [100], [110], [111]\}$;
- $A_3 = \{[001], [100], [101]\}$.

Solution: It is easily verified that A_1 is a subspace, since it contains the all-zero vector and the sum of any two tuples in A_1 is also in A_1. A_2 is not a subspace because it is not closed under addition, as $110 + 111 = 001 \notin A_2$. A_3 is not a subspace because it is not closed under addition $(001 + 001 = 000 \notin A_3)$ and it does not contain the all-zero vector.

Intuitively, the greater the distance between codewords in a given code, the less chance that errors introduced by the channel will cause a transmitted codeword to be decoded as a different codeword. We define the *Hamming distance* between two codewords $\mathbf{C}_i$ and $\mathbf{C}_j$, denoted as $d(\mathbf{C}_i, \mathbf{C}_j)$ or d_{ij}, as the number of elements in which they differ:

$$d_{ij} = \sum_{l=1}^{n} (\mathbf{C}_i(l) + \mathbf{C}_j(l)), \tag{8.4}$$

where $\mathbf{C}_m(l)$ denotes the lth bit in $\mathbf{C}_m$. For example, if $\mathbf{C}_i = [00101]$ and $\mathbf{C}_j = [10011]$ then $d_{ij} = 3$. We define the *weight* of a given codeword $\mathbf{C}_i$ as the number of 1-bits in the codeword, so $\mathbf{C}_i = [00101]$ has weight 2. The weight of a given codeword $\mathbf{C}_i$ is just its Hamming distance d_{0i} from the all-zero codeword $\mathbf{C}_0 = [00\ldots0]$ or, equivalently, the sum of its elements:

$$w(\mathbf{C}_i) = \sum_{l=1}^{n} \mathbf{C}_i(l). \tag{8.5}$$

Since $0+0=1+1=0$, the Hamming distance between $\mathbf{C}_i$ and $\mathbf{C}_j$ is equal to the weight of $\mathbf{C}_i+\mathbf{C}_j$. For example, with $\mathbf{C}_i=[00101]$ and $\mathbf{C}_j=[10011]$ as given before, $w(\mathbf{C}_i)=2, w(\mathbf{C}_j)=3$, and $d_{ij}=w(\mathbf{C}_i+\mathbf{C}_j)=w([10110])=3$. Since the Hamming distance between any two codewords equals the weight of their sum and since this sum is also a codeword, we can determine the minimum distance between all codewords in a code by just looking at the minimum distance between all nonzero codewords and the all-zero codeword $\mathbf{C}_0$. Thus, we define the minimum distance of a code as

$$d_{\min}=\min_{i,i\neq 0} d_{0i}. \tag{8.6}$$

A detailed analysis of block code error probability under soft decision decoding is given in Section 8.2.7. In particular, the probability of mistaking the all-zero codeword $\mathbf{C}_0$ for another codeword $\mathbf{C}_i$ at distance d_{0i} under BPSK transmission over an AWGN channel is derived in (8.52) as $P_e(\mathbf{C}_i)=Q(\sqrt{2d_{0i}\gamma_b R_c})$. Since $d_{\min}\leq d_{0i}$ and there are a total of 2^k codewords in an (n,k) linear block code, we thus have from the union bound that the codeword error probability satisfies

$$P_e\leq (2^k-1)Q\left(\sqrt{2d_{\min}\gamma_b R_c}\right). \tag{8.7}$$

Since each of the codewords conveys k bits, the bit error probability is bounded as

$$P_b\leq \frac{2^k-1}{k}Q\left(\sqrt{2d_{\min}\gamma_b R_c}\right). \tag{8.8}$$

At high SNRs the codeword error probability is dominated by the probability of mistaking a codeword for one of its N nearest neighbors at distance $d_{\min}$. Hence at high SNRs we often use the approximation

$$P_b\approx \frac{N}{k}Q\left(\sqrt{2d_{\min}\gamma_b R_c}\right). \tag{8.9}$$

Recall that uncoded BPSK has error probability $P_b=Q(\sqrt{2\gamma_b})$. Comparing this expression to the bound (8.8) we see that the SNR is increased by the code's minimum distance $d_{\min}$, but this SNR must also be distributed over the n coded bits rather than just the k information bits, hence SNR is reduced by $R_c=k/n$. Thus, coding provides a reduction in P_b relative to uncoded transmission when $d_{\min}R_c>1$ and $Q(\sqrt{2d_{\min}\gamma_b R_c})\ll(2^k-1)/k$. To illustrate this behavior, in Figure 8.3 we plot the P_b approximation (8.9) for the $(32,6,16)$ biorthogonal block code, which has minimum distance $d_{\min}=16$ and $N=16\approx 2^{k-1}$ nearest neighbors. Also shown is P_b for uncoded BPSK and the Shannon limit of -1.59 dB as $\eta=R/B\to\infty$. We see in the figure a large coding gain at high SNR which shrinks as SNR decreases, becoming negative at around $\gamma_b=0$ dB. We also see that the performance of this code is far from the Shannon limit at practical P_b values.

8.2.2 Generator Matrix

The generator matrix is a compact description of how codewords are generated from information bits in a linear block code. The design goal in linear block codes is to find generator matrices such that their corresponding codes are easy to encode and decode yet have powerful error correction/detection capabilities. Consider an (n,k) code with k information bits, denoted as

$$\mathbf{U}_i=[u_{i1},\ldots,u_{ik}], \tag{8.10}$$

that are encoded into the codeword

$$\mathbf{C}_i=[c_{i1},\ldots,c_{in}]. \tag{8.11}$$

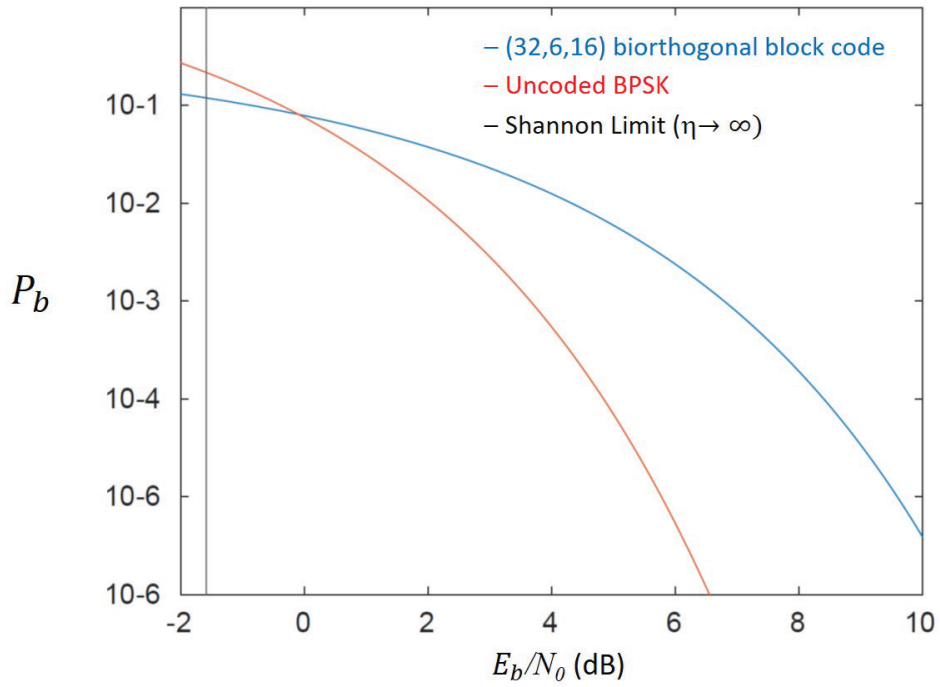


Figure 8.3: P_b versus γ_b for the $(32, 6, 16)$ biorthogonal block code, which has minimum distance 16 and $N = 62 \approx 2^{k-1}$ nearest neighbors. Also shown is P_b for uncoded BPSK and the Shannon limit of -1.59 dB as $\eta \to \infty$.

We represent the encoding operation as a set of n equations defined by

$$c_{ij} = u_{i1}g_{1j} + u_{i2}g_{2j} + \cdots + u_{ik}g_{kj}, \quad j = 1, \ldots, n, \tag{8.12}$$

where g_{ij} is binary (0 or 1) and where binary (standard) multiplication is used. We can write these n equations in matrix form as

$$\mathbf{C}_i = \mathbf{U}_i\mathbf{G}, \tag{8.13}$$

where the $k \times n$ *generator matrix* $\mathbf{G}$ for the code is defined as

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}. \tag{8.14}$$

If we denote the lth row of $\mathbf{G}$ as $\mathbf{g}_l = [g_{l1}, \ldots, g_{ln}]$, then we can write any codeword $\mathbf{C}_i$ as linear combinations of these row vectors as follows:

$$\mathbf{C}_i = u_{i1}\mathbf{g}_1 + u_{i2}\mathbf{g}_2 + \cdots + u_{ik}\mathbf{g}_k. \tag{8.15}$$

Since a linear (n, k) block code is a subspace of dimension k in the larger n-dimensional space, it follows that the k row vectors $\{\mathbf{g}_l\}_{l=1}^{k}$ of $\mathbf{G}$ must be linearly independent so that they span the k-dimensional subspace associated with the 2^k codewords. Hence, $\mathbf{G}$ has rank k. Since the set of basis vectors for this subspace is not unique, the generator matrix is also not unique.

A *systematic* linear block code is described by a generator matrix of the form

$$\mathbf{G} = [\mathbf{I}_k|\mathbf{P}] = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ 0 & 1 & \cdots & 0 & p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{k1} & p_{k2} & \cdots & p_{k(n-k)} \end{array}\right], \tag{8.16}$$

where $\mathbf{I}_k$ is the $k \times k$ identity matrix and $\mathbf{P}$ is a $k \times (n-k)$ matrix that determines the redundant, or parity, bits to be used for error correction or detection. The codeword output from a systematic encoder is of the form

$$\mathbf{C}_i = \mathbf{U}_i\mathbf{G} = \mathbf{U}_i[\mathbf{I}_k|\mathbf{P}] = [u_{i1}, \ldots, u_{ik}, p_1, \ldots, p_{(n-k)}], \tag{8.17}$$

where the first k bits of the codeword are the original information bits and the last $(n-k)$ bits of the codeword are the parity bits obtained from the information bits as

$$p_j = u_{i1}p_{1j} + \cdots + u_{ik}p_{kj}, \quad j = 1, \ldots, n-k. \tag{8.18}$$

Note that any generator matrix for an (n, k) linear block code can be reduced by row operations and column permutations to a generator matrix in systematic form.

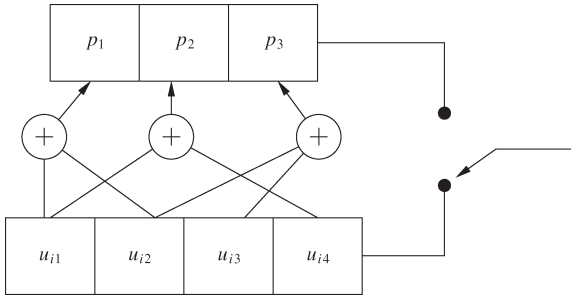


Figure 8.4: Implementation of (7,4) binary code.

Example 8.2: Systematic linear block codes are typically implemented with $n-k$ modulo-2 adders tied to the appropriate stages of a shift register. The resulting parity bits are appended to the end of the information bits to form the codeword. Find the corresponding implementation for generating a (7,4) binary code with the generator matrix

$$\mathbf{G} = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right]. \tag{8.19}$$

Solution: The matrix $\mathbf{G}$ is already in systematic form with

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{8.20}$$

Let P_{lj} denote the ljth element of $\mathbf{P}$. By (8.18) we see that the first parity bit in the codeword is $p_1 = u_{i1}P_{11} + u_{i2}P_{21} + u_{i3}P_{31} + u_{i4}P_{41} = u_{i1} + u_{i2}$. Similarly, the second parity bit is $p_2 = u_{i1}P_{12} + u_{i2}P_{22} + u_{i3}P_{32} + u_{i4}P_{42} = u_{i1} + u_{i4}$ and the third parity bit is $p_3 = u_{i1}P_{13} + u_{i2}P_{23} + u_{i3}P_{33} + u_{i4}P_{43} = u_{i2} + u_{i3}$. The shift register implementation to generate these parity bits is shown in Figure 8.4. The codeword output is $[u_{i1}u_{i2}u_{i3}u_{i4}p_1p_2p_3]$, where the switch is in the down position to output the systematic bits u_{ij}, $j = 1, \ldots, 4$, of the code or in the up position to output the parity bits p_j, $j = 1, 2, 3$, of the code.

8.2.3 Parity-Check Matrix and Syndrome Testing

The parity-check matrix is used to decode linear block codes with generator matrix $\mathbf{G}$. The parity-check matrix $\mathbf{H}$ corresponding to a generator matrix $\mathbf{G} = [\mathbf{I}_k|\mathbf{P}]$ is defined as

$$\mathbf{H} = [\mathbf{P}^T|\mathbf{I}_{n-k}]. \tag{8.21}$$

It is easily verified that $\mathbf{G}\mathbf{H}^T = \mathbf{0}_{k,n-k}$, where $\mathbf{0}_{k,n-k}$ denotes an all-zero $k \times (n-k)$ matrix. Recall that a given codeword $\mathbf{C}_i$ in the code is obtained by multiplication of the information bit sequence $\mathbf{U}_i$ by the generator matrix $\mathbf{G}$: $\mathbf{C}_i = \mathbf{U}_i\mathbf{G}$. Thus,

$$\mathbf{C}_i\mathbf{H}^T = \mathbf{U}_i\mathbf{G}\mathbf{H}^T = \mathbf{0}_{n-k} \tag{8.22}$$

for any input sequence $\mathbf{U}_i$, where $\mathbf{0}_{n-k}$ denotes the all-zero row vector of length $n-k$. Thus, multiplication of any valid codeword with the parity-check matrix results in an all-zero vector. This property is used to determine whether the received vector is a valid codeword or has been corrupted, based on the notion of *syndrome testing*, which we now define.

Let $\mathbf{R}$ be the received codeword resulting from transmission of codeword $\mathbf{C}$. In the absence of channel errors, $\mathbf{R} = \mathbf{C}$. However, if the transmission is corrupted, then one or more of the codeword symbols in $\mathbf{R}$ will differ from those in $\mathbf{C}$. We therefore write the received codeword as

$$\mathbf{R} = \mathbf{C} + \mathbf{e}, \tag{8.23}$$

where $\mathbf{e} = [e_1, e_2, \ldots, e_n]$ is the *error pattern* indicating which codeword symbols were corrupted by the channel. We define the *syndrome* of $\mathbf{R}$ as

$$\mathbf{S} = \mathbf{R}\mathbf{H}^T. \tag{8.24}$$

If $\mathbf{R}$ is a valid codeword (i.e., $\mathbf{R} = \mathbf{C}_i$ for some i) then $\mathbf{S} = \mathbf{C}_i\mathbf{H}^T = \mathbf{0}_{n-k}$ by (8.22). Thus, the syndrome equals the all-zero vector if the transmitted codeword is not corrupted – or is corrupted in a manner such that the received codeword is a valid codeword in the code but is different from the transmitted codeword, resulting in an undetectable error. If the received codeword $\mathbf{R}$ contains detectable errors, then $\mathbf{S} \neq \mathbf{0}_{n-k}$. If the received codeword contains correctable errors, the syndrome identifies the error pattern corrupting the transmitted codeword, and these

errors can then be corrected. Note that the syndrome is a function only of the error pattern $\mathbf{e}$ and not the transmitted codeword $\mathbf{C}$, since

$$\mathbf{S} = \mathbf{R}\mathbf{H}^T = (\mathbf{C} + \mathbf{e})\mathbf{H}^T = \mathbf{C}\mathbf{H}^T + \mathbf{e}\mathbf{H}^T = \mathbf{0}_{n-k} + \mathbf{e}\mathbf{H}^T. \tag{8.25}$$

Because $\mathbf{S} = \mathbf{e}\mathbf{H}^T$ corresponds to $n - k$ equations in n unknowns, there are 2^k possible error patterns that can produce a given syndrome $\mathbf{S}$. However, since the probability of bit error is typically small and independent for each bit, the most likely error pattern is the one with minimal weight, corresponding to the least number of errors introduced in the channel. Thus, if an error pattern $\hat{\mathbf{e}}$ is the most likely error associated with a given syndrome $\mathbf{S}$, the transmitted codeword is typically decoded as

$$\hat{\mathbf{C}} = \mathbf{R} + \hat{\mathbf{e}} = \mathbf{C} + \mathbf{e} + \hat{\mathbf{e}}. \tag{8.26}$$

If the most likely error pattern does occur, then $\hat{\mathbf{e}} = \mathbf{e}$ and $\hat{\mathbf{C}} = \mathbf{C}$ – that is, the corrupted codeword is correctly decoded. The decoding process and associated error probability will be covered in Section 8.2.6

Let $\mathbf{C}_w$ denote a codeword in a given (n, k) code with minimum weight (excluding the all-zero codeword). Then $\mathbf{C}_w\mathbf{H}^T = \mathbf{0}_{n-k}$ is just the sum of $d_{\min}$ columns of $\mathbf{H}^T$, since $d_{\min}$ equals the number of l-bits (the weight) in the minimum weight codeword of the code. Since the rank of $\mathbf{H}^T$ is at most $n - k$, this implies that the minimum distance of an (n, k) block code is upper bounded by

$$d_{\min} \leq n - k + 1, \tag{8.27}$$

which is referred to as the *Singleton bound*.

8.2.4 Cyclic Codes

Cyclic codes are a subclass of linear block codes in which all codewords in a given code are cyclic shifts of one another. Specifically, if the codeword $\mathbf{C} = [c_0, c_1, \ldots, c_{n-1}]$ is a codeword in a given code, then a cyclic shift by 1, denoted as $\mathbf{C}^{(1)} = [c_{n-1}, c_0, \ldots, c_{n-2}]$, is also a codeword. More generally, any cyclic shift $C^{(i)} = [c_{n-i}, c_{n-i+1}, \ldots, c_{n-i-1}]$ is also a codeword. The cyclic nature of cyclic codes creates a nice structure that allows their encoding and decoding functions to be of much lower complexity than the matrix multiplications associated with encoding and decoding for general linear block codes. Thus, most linear block codes used in practice are cyclic codes.

Cyclic codes are generated via a *generator polynomial* instead of a generator matrix. The generator polynomial $g(X)$ for an (n, k) cyclic code has degree $n - k$ and is of the form

$$g(X) = g_0 + g_1X + \cdots + g_{n-k}X^{n-k}, \tag{8.28}$$

where g_i is binary (0 or 1) and $g_0 = g_{n-k} = 1$. The k-bit information sequence $[u_0, \ldots, u_{k-1}]$ is also written in polynomial form as the *message polynomial*

$$u(X) = u_0 + u_1X + \cdots + u_{k-1}X^{k-1}. \tag{8.29}$$

The codeword associated with a given k-bit information sequence is obtained from the polynomial coefficients of the generator polynomial multiplied by the message polynomial; thus, the codeword $C = [c_0, \ldots, c_{n-1}]$ is obtained from

$$c(X) = u(X)g(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}. \tag{8.30}$$

A codeword described by a polynomial $c(X)$ is a valid codeword for a cyclic code with generator polynomial $g(X)$ if and only if $g(X)$ divides $c(X)$ with no remainder (no remainder polynomial terms) – that is, if and only if

$$\frac{c(X)}{g(X)} = q(X) \tag{8.31}$$

for a polynomial $q(X)$ of degree less than k.

Example 8.3: Consider a (7, 4) cyclic code with generator polynomial $g(X) = 1 + X^2 + X^3$. Determine if the codewords described by polynomials $c_1(X) = 1 + X^2 + X^5 + X^6$ and $c_2(X) = 1 + X^2 + X^3 + X^5 + X^6$ are valid codewords for this generator polynomial.

Solution: Division of binary polynomials is similar to division of standard polynomials except that, under binary addition, subtraction is the same as addition. Dividing $c_1(X) = 1 + X^2 + X^5 + X^6$ by $g(X) = 1 + X^2 + X^3$, we have:

$$\begin{array}{r} X^3 + 1 \\ \hline X^3 + X^2 + 1 \;\big)X^6 + X^5 + \;X^2 + 1 \\ \quad X^6 + X^5 + \;X^3 \\ \hline \qquad\qquad\quad X^3 + X^2 + 1 \\ \qquad\qquad\quad X^3 + X^2 + 1 \\ \hline \qquad\qquad\quad 0. \end{array} \tag{8.32}$$

Since $g(X)$ divides $c(X)$ with no remainder, it is a valid codeword. In fact, we have $c_1(X) = (1 + X^3)g(X) = u(X)g(X)$ and so the information bit sequence corresponding to $c_1(X)$ is $\mathbf{U} = [1001]$, corresponding to the coefficients of the message polynomial $u(X) = 1 + X^3$.

Dividing $c_2(X) = 1 + X^2 + X^3 + X^5 + X^6$ by $g(X) = 1 + X^2 + X^3$ yields

$$\begin{array}{r} X^3 + 1 \\ \hline X^3 + X^2 + 1 \;\big)X^6 + X^5 + X^3 + X^2 + 1 \\ \quad X^6 + X^5 + X^3 \\ \hline \qquad\qquad\quad X^2 + 1, \end{array} \tag{8.33}$$

where we note that there is a remainder of $X^2 + 1$ in the division. Thus, $c_2(X)$ is not a valid codeword for the code corresponding to this generator polynomial.

Recall that systematic linear block codes have the first k codeword symbols equal to the information bits and the remaining codeword symbols equal to the parity bits. A cyclic code can be put in systematic form by first multiplying the message polynomial $u(X)$ by X^{n-k}, yielding

$$X^{n-k}u(X) = u_0X^{n-k} + u_1X^{n-k+1} + \cdots + u_{k-1}X^{n-1}. \tag{8.34}$$

This shifts the message bits to the k rightmost digits of the codeword polynomial. If we next divide (8.34) by $g(X)$, we obtain

$$\frac{X^{n-k}u(X)}{g(X)} = q(X) + \frac{p(X)}{g(X)}, \tag{8.35}$$

where $q(X)$ is a polynomial of degree at most $k-1$ and $p(X)$ is a remainder polynomial of degree at most $n-k-1$. Multiplying (8.35) through by $g(X)$, we now have

$$X^{n-k}u(X) = q(X)g(X) + p(X). \tag{8.36}$$

Adding $p(X)$ to both sides yields

$$p(X) + X^{n-k}u(X) = q(X)g(X). \tag{8.37}$$

This implies that $p(X) + X^{n-k}u(X)$ is a valid codeword, since it is divisible by $g(X)$ with no remainder. The codeword is described by the n coefficients of the codeword polynomial $p(X) + X^{n-k}u(X)$. Note that we can express $p(X)$ (of degree $n-k-1$) as

$$p(X) = p_0 + p_1X + \cdots + p_{n-k-1}X^{n-k-1}. \tag{8.38}$$

Combining (8.34) and (8.38), we get

$$\begin{aligned} &p(X) + X^{n-k}u(X) \\ &\quad = p_0 + p_1X + \cdots + p_{n-k-1}X^{n-k-1} + u_0X^{n-k} + u_1X^{n-k+1} + \cdots + u_{k-1}X^{n-1}. \end{aligned} \tag{8.39}$$

Thus, the codeword corresponding to this polynomial has the first k bits consisting of the message bits $[u_0, \ldots, u_k]$ and the last $n-k$ bits consisting of the parity bits $[p_0, \ldots, p_{n-k-1}]$, as is required for the systematic form.

Note that the systematic codeword polynomial is generated in three steps: first multiplying the message polynomial $u(X)$ by X^{n-k}; then dividing $X^{n-k}u(X)$ by $g(X)$ to obtain the remainder polynomial $p(X)$ (along with the quotient polynomial $q(X)$, which is not used); and finally adding $p(X)$ to $X^{n-k}u(X)$ to get (8.39). The polynomial multiplications are straightforward to implement, and the polynomial division is easily implemented with a feedback shift register (see [6, Chap. 8.1] and [7, Chap. 6.7]). Thus, codeword generation for systematic cyclic codes has very low cost and low complexity.

Let us now consider how to characterize channel errors for cyclic codes. The codeword polynomial corresponding to a transmitted codeword is of the form

$$c(X) = u(X)g(X). \tag{8.40}$$

The received codeword can also be written in polynomial form as

$$r(X) = c(X) + e(X) = u(X)g(X) + e(X), \tag{8.41}$$

where $e(X)$ is the error polynomial with coefficients equal to 1 where errors occur. For example, if the transmitted codeword is $\mathbf{C} = [1011001]$ and the received codeword is $\mathbf{R} = [1111000]$, then $e(X) = X + X^{n-1}$. The *syndrome polynomial* $s(X)$ for the received codeword is defined as the remainder when $r(X)$ is divided by $g(X)$, so $s(X)$ has degree $n-k-1$. But by (8.41), the syndrome polynomial $s(X)$ is equivalent to the error polynomial $e(X)$ modulo $g(X)$. Moreover, we obtain the syndrome through a division circuit similar to the one used for generating the code. As stated previously, this division circuit is typically implemented using a feedback shift register, resulting in a low-cost implementation of low complexity.

8.2.5 Hard Decision Decoding (HDD)

The probability of error for linear block codes depends on whether the decoder uses soft decisions or hard decisions. In hard decision decoding (HDD), each coded bit is demodulated as a 0 or 1 – that is, the demodulator detects each coded bit (symbol) individually. For example, in BPSK, the received symbol is decoded as a 1 if it is closer to $\sqrt{E_b}$ and as a 0 if it is closer to $-\sqrt{E_b}$. This form of demodulation removes information that can be used by the channel decoder. In particular, for the BPSK example the distance of the received bit from $\sqrt{E_b}$ and $-\sqrt{E_b}$ can be used in the channel decoder to make better decisions about the transmitted codeword. In soft decision decoding, these distances are used in the decoding process. Soft decision decoding of linear block codes, which is more common in wireless systems than hard decision decoding, is treated in Section 8.2.7.

Hard decision decoding uses *minimum distance* decoding based on Hamming distance. In minimum distance decoding the n bits corresponding to a codeword are first demodulated to a 0 or 1, and the demodulator output is then passed to the decoder. The decoder compares this received codeword to the 2^k possible codewords that constitute the code and decides in favor of the codeword that is closest in Hamming distance to (differs in the least number of bits from) the received codeword. Mathematically, for a received codeword $\mathbf{R}$, the decoder uses the formula

$$\text{pick } \mathbf{C}_j \text{ s.t. } d(\mathbf{C}_j, \mathbf{R}) \leq d(\mathbf{C}_i, \mathbf{R}),\ \forall i \neq j. \tag{8.42}$$

If there is more than one codeword with the same minimum distance to $\mathbf{R}$, one of these is chosen at random by the decoder.

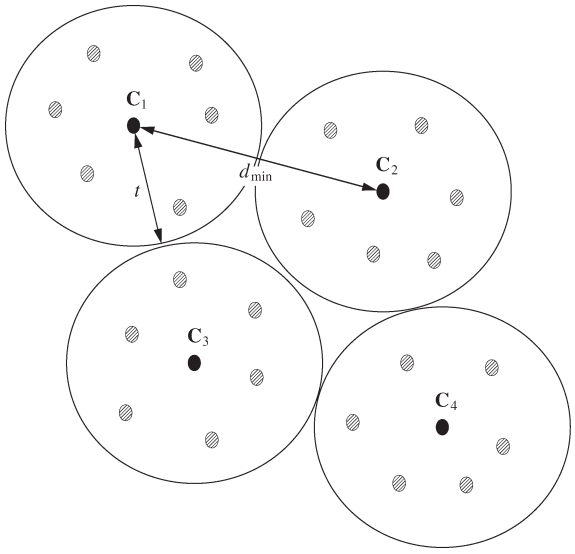


Figure 8.5: Maximum likelihood decoding in code space.

Maximum likelihood decoding picks the transmitted codeword that has the highest probability of having produced the received codeword. In other words, given the received codeword $\mathbf{R}$, the maximum likelihood decoder chooses the codeword $\mathbf{C}_j$, where j is selected as

$$j = \arg\max_i p(\mathbf{R}|\mathbf{C}_i),\ \ i = 0, \ldots, 2^k - 1. \tag{8.43}$$

Since the most probable error event in an AWGN channel is the event with the minimum number of errors needed to produce the received codeword, the minimum distance criterion (8.42) and the maximum likelihood criterion

(8.43) are equivalent. Once the maximum likelihood codeword $\mathbf{C}_i$ is determined, it is decoded to the k bits that produce codeword $\mathbf{C}_i$.

Because maximum likelihood detection of codewords is based on a distance decoding metric, we can best illustrate this process in code space, as shown in Figure 8.5. The minimum Hamming distance between codewords, which are illustrated by the black dots in this figure, is $d_{\min}$. Each codeword is centered inside a sphere of radius $t = \lfloor .5d_{\min} \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. The shaded dots represent received sequences where one or more bits differ from those of the transmitted codeword. The figure indicates the Hamming distance between $\mathbf{C}_1$ and $\mathbf{C}_2$.

Minimum distance decoding can be used either to detect or to correct errors. Detected errors in a data block cause either the data to be dropped or a retransmission of the data. Error correction allows the corruption in the data to be reversed. For error correction the minimum distance decoding process ensures that a received codeword lying within a Hamming distance t from the transmitted codeword will be decoded correctly. Thus, the decoder can correct up to t errors, as can be seen from Figure 8.5: since received codewords corresponding to t or fewer errors will lie within the sphere centered around the correct codeword, it will be decoded as that codeword using minimum distance decoding. We see from Figure 8.5 that the decoder can detect all error patterns of $d_{\min} - 1$ errors. In fact, a decoder for an (n, k) code can detect $2^n - 2^k$ possible error patterns. The reason is that there are $2^k - 1$ nondetectable errors, corresponding to the case where a corrupted codeword is exactly equal to a codeword in the set of possible codewords (of size 2^k) that is not equal to the transmitted codeword. Since there are $2^n - 1$ total possible error patterns, this yields $2^n - 2^k$ detectable error patterns. Note that this is not hard decision decoding because we are not correcting errors, just detecting them.

Example 8.4: A $(5, 2)$ code has codewords $\mathbf{C}_0 = [00000]$, $\mathbf{C}_1 = [01011]$, $\mathbf{C}_2 = [10101]$, and $\mathbf{C}_3 = [11110]$. Suppose the all-zero codeword $\mathbf{C}_0$ is transmitted. Find the set of error patterns corresponding to nondetectable errors for this codeword transmission.

Solution: The nondetectable error patterns correspond to the three nonzero codewords. That is, $\mathbf{e}_1 = [01011]$, $\mathbf{e}_2 = [10101]$, and $\mathbf{e}_3 = [11110]$ are nondetectable error patterns, since adding any of these to $\mathbf{C}_0$ results in a valid codeword.

8.2.6 Probability of Error for HDD in AWGN

The probability of codeword error, P_e, is defined as the probability that a transmitted codeword is decoded in error. Under hard decision decoding a received codeword *may* be decoded in error if it contains more than t errors (it will not be decoded in error if there is not an alternative codeword closer to the received codeword than the transmitted codeword). The error probability is thus bounded above by the probability that more than t errors occur. Since the bit errors in a codeword occur independently on an AWGN channel, this probability is given by

$$P_e \leq \sum_{j=t+1}^{n} \binom{n}{j} p^j (1-p)^{n-j}, \tag{8.44}$$

where p is the probability of error associated with transmission of the bits in the codeword. Thus, p corresponds to the error probability associated with uncoded modulation for the given energy per codeword symbol, as treated in Chapter 6 for AWGN channels. For example, if the codeword symbols are sent via BPSK modulation then $p = Q(\sqrt{2E_c/N_0})$, where E_c is the energy per codeword symbol and $N_0/2$ is the noise power spectral density. Since there are k/n information bits per codeword symbol, the relationship between the energy per bit and the energy per symbol is $E_c = kE_b/n$. Thus, powerful block codes with a large number of parity bits (k/n small) reduce the channel energy per symbol and therefore increase the error probability in demodulating the codeword

symbols. However, the error correction capability of these codes typically more than compensates for this reduction, especially at high SNRs. At low SNRs this may not happen, in which case the code exhibits *negative coding gain* − it has a higher error probability than uncoded modulation. The bound (8.44) holds with equality when the decoder corrects exactly t or fewer errors in a codeword and cannot correct for more than t errors in a codeword. A code with this property is called a *perfect* code.

At high SNRs, the most likely way to make a codeword error is to mistake a codeword for one of its nearest neighbors. Nearest neighbor errors yield a pair of upper and lower bounds on error probability. The lower bound is the probability of mistaking a codeword for a given nearest neighbor at distance $d_{\min}$:

$$P_e \geq \sum_{j=t+1}^{d_{\min}} \binom{d_{\min}}{j} p^j (1-p)^{d_{\min}-j}. \tag{8.45}$$

The upper bound, a union bound, assumes that all of the other $2^k - 1$ codewords are at distance $d_{\min}$ from the transmitted codeword. Thus, the union bound is just $2^k - 1$ times (8.45), the probability of mistaking a given codeword for a nearest neighbor at distance $d_{\min}$:

$$P_e \leq (2^k - 1) \sum_{j=t+1}^{d_{\min}} \binom{d_{\min}}{j} p^j (1-p)^{d_{\min}-j}. \tag{8.46}$$

When the number of codewords is large or the SNR is low, both of these bounds are quite loose.

A tighter upper bound can be obtained by applying the Chernoff bound, $P(X \geq x) \leq e^{-x^2/2}$ for X a zero-mean, unit-variance, Gaussian random variable, to compute codeword error probability. Using this bound, it can be shown [8, Chap. 5.2] that the probability of decoding the all-zero codeword as the jth codeword with weight w_j is upper bounded by

$$P(w_j) \leq [4p(1-p)]^{w_j/2}. \tag{8.47}$$

Since the probability of decoding error is upper bounded by the probability of mistaking the all-zero codeword for any of the other codewords, we obtain the upper bound

$$P_e \leq \sum_{j=1}^{2^k-1} [4p(1-p)]^{w_j/2}. \tag{8.48}$$

This bound requires the weight distribution $\{w_j\}_{j=1}^{2^k-1}$ for all codewords (other than the all-zero codeword corresponding to $j = 0$) in the code. A simpler, slightly looser upper bound is obtained from (8.48) by using $d_{\min}$ instead of the individual codeword weights. This simplification yields the bound

$$P_e \leq (2^k - 1)[4p(1-p)]^{d_{\min}/2}. \tag{8.49}$$

Note that the probability of codeword error P_e depends on p, which is a function of the Euclidean distance between modulation points associated with the transmitted codeword symbols. In fact, the best codes for AWGN channels should not be based on Hamming distance: they should be based on maximizing the Euclidean distance between the codewords after modulation. However, this requires that the channel code be designed jointly with the modulation. This is the basic concept of coded modulation, which will be discussed in Section 8.9. However, Hamming distance is a better measure of code performance in fading when codes are combined with interleaving, as discussed in Section 8.10.

The probability of bit error after decoding the received codeword depends in general on the specific code and decoder and in particular on how bits are mapped to codewords, as in the bit mapping procedure associated with nonbinary modulation. This bit error probability is often approximated as

$$P_b \approx \frac{1}{n} \sum_{j=t+1}^{n} j \binom{n}{j} p^j (1-p)^{n-j} \tag{8.50}$$

(see [7, Chap. 6.5]); for $t = 1$, this can be simplified [7] to $P_b \approx p - p(1-p)^{n-1}$.

Example 8.5: Consider a $(24, 12)$ linear block code with a minimum distance $d_{\min} = 8$ (an extended Golay code, discussed in Section 8.2.8, is one such code). Find P_e based on the loose bound (8.49), assuming the codeword symbols are transmitted over the channel using BPSK modulation with $E_b/N_0 = 10\,\mathrm{dB}$. Also find P_b for this code using the approximation $P_b = P_e/k$ and compare with the bit error probability for uncoded modulation.
Solution: For $E_b/N_0 = 10\,\mathrm{dB}$ we have $E_c/N_0 = \frac{12}{24}10 = 5$. Thus, $p = Q(\sqrt{10}) = 7.82 \cdot 10^{-4}$. Using this value in (8.49) with $k = 12$ and $d_{\min} = 8$ yields $P_e \leq 3.92 \cdot 10^{-7}$. Using the P_b approximation we get $P_b \approx (1/k)P_e = 3.27 \cdot 10^{-8}$. For uncoded modulation we have $P_b = Q(\sqrt{2E_b/N_0}) = Q(\sqrt{20}) = 3.87 \cdot 10^{-6}$. Thus we obtain over two orders of magnitude performance gain with this code. Note that the loose bound can be orders of magnitude away from the true error probability, so this calculation may significantly underestimate the performance gain of the code.

8.2.7 Probability of Error for SDD in AWGN

The HDD described in the previous section discards information that can reduce probability of codeword error. For example, in BPSK, the transmitted signal constellation is $\pm\sqrt{E_b}$ and the received symbol after matched filtering is decoded as a 1 if it is closer to $\sqrt{E_b}$ and as a 0 if it is closer to $-\sqrt{E_b}$. Thus, the distance of the received symbol from $\sqrt{E_b}$ and $-\sqrt{E_b}$ is not used in decoding, yet this information can be used to make better decisions about the transmitted codeword. When these distances are used in the channel decoder it is called *soft decision decoding* (SDD), since the demodulator does not make a hard decision about whether a 0 or 1 bit was transmitted but rather makes a soft decision corresponding to the distance between the received symbol and the symbol corresponding to a 0-bit or a 1-bit transmission. We now describe the basic premise of SDD for BPSK modulation; these ideas are easily extended to higher-level modulations.

Consider a codeword transmitted over a channel using BPSK. As in the case of HDD, the energy per codeword symbol is $E_c = (k/n)E_b$. If the jth codeword symbol is a 1 it will be received as $r_j = \sqrt{E_c} + n_j$ and if it is a 0 it will be received as $r_j = -\sqrt{E_c} + n_j$, where n_j is the AWGN sample of mean zero and variance $N_0/2$ associated with the receiver. In SDD, given a received codeword $\mathbf{R} = [r_1, \ldots, r_n]$, the decoder forms a *correlation metric* $C(\mathbf{R},\ \mathbf{C}_i)$ for each codeword $\mathbf{C}_i$, $i = 0, \ldots, 2^k - 1$, in the code and then the decoder chooses the codeword $\mathbf{C}_i$ with the highest correlation metric. The correlation metric is defined as

$$C(\mathbf{R},\ \mathbf{C}_i) = \sum_{j=1}^{n} (2c_{ij} - 1) r_j, \tag{8.51}$$

where c_{ij} denotes the jth coded bit in the codeword $\mathbf{C}_i$. If $c_{ij} = 1$ then $2c_{ij} - 1 = 1$ and if $c_{ij} = 0$ then $2c_{ij} - 1 = -1$. Thus the received codeword symbol is weighted by the polarity associated with the corresponding symbol in the codeword for which the correlation metric is being computed. Hence $C(\mathbf{R}, \mathbf{C}_i)$ is large when most of the received symbols have a large magnitude and the same polarity as the corresponding symbols in $\mathbf{C}_i$, is smaller when most of the received symbols have a small magnitude and the same polarity as the corresponding

symbols in $\mathbf{C}_i$, and is typically negative when most of the received symbols have a different polarity than the corresponding symbols in $\mathbf{C}_i$. In particular, at very high SNRs, if $\mathbf{C}_i$ is transmitted then $C(\mathbf{R}, \mathbf{C}_i) \approx n\sqrt{E_c}$ while $C(\mathbf{R}, \mathbf{C}_j) < n\sqrt{E_c}$ for $j \neq i$.

For an AWGN channel, the probability of codeword error is the same for any codeword of a linear code. Error analysis is typically easiest when assuming transmission of the all-zero codeword. Let us therefore assume that the all-zero codeword $\mathbf{C}_0$ is transmitted using BPSK, and the corresponding received codeword is $\mathbf{R}$. To correctly decode $\mathbf{R}$, we must have that $C(\mathbf{R},\ \mathrm{C}_0) > C(\mathbf{R},\ \mathbf{C}_i)$, $i = 1, \ldots, 2^k - 1$. Let w_i denote the Hamming weight of the ith codeword $\mathbf{C}_i$, which equals the number of 1-bits in $\mathbf{C}_i$. Then, conditioned on the transmitted codeword $\mathbf{C}_i$, it follows that $C(\mathbf{R},\ \mathbf{C}_i)$ is Gauss distributed with mean $\sqrt{E_c}n(1 - 2w_i/n)$ and variance $nN_0/2$. Note that the correlation metrics are not independent, since they are all functions of $\mathbf{R}$. The probability $P_e(\mathbf{C}_i) = p(C(\mathbf{R},\ \mathbf{C}_0) < C(\mathbf{R},\ \mathbf{C}_i))$ can be shown to equal the probability that a Gauss-distributed random variable with variance $2w_iN_0$ is less than $-2w_i\sqrt{E_c}$ (see Problem 8-7); that is,

$$P_e(\mathbf{C}_i) = Q\left(\frac{2w_i\sqrt{E_c}}{\sqrt{2w_iN_0}}\right) = Q(\sqrt{2w_i\gamma_bR_c}). \tag{8.52}$$

Then, by the union bound, the probability of error is upper bounded by the sum of pairwise error probabilities relative to each $\mathbf{C}_i$:

$$P_e \leq \sum_{i=1}^{2^k-1} P_e(\mathbf{C}_i) = \sum_{i=1}^{2^k-1} Q(\sqrt{2w_i\gamma_bR_c}). \tag{8.53}$$

Computing (8.53) requires the weight distribution w_i, $i = 1, \ldots, 2^k - 1$, of the code. This bound can be simplified by noting that $w_i \geq d_{\min}$, so

$$P_e \leq (2^k - 1)Q(\sqrt{2d_{\min}\gamma_bR_c}). \tag{8.54}$$

At high SNRs γ_b is large, so $Q(\sqrt{2w_i\gamma_bR_c}) \ll Q(\sqrt{2d_{\min}\gamma_bR_c})$ for $w_i < d_{\min}$ given the exponential decrease of the Q function. Hence, at high SNRs, (8.53) is well-approximated by the error associated with mistaking a codeword for one of its N nearest neighbors at distance $d_{\min}$, as follows:

$$P_e \approx NQ(\sqrt{2d_{\min}\gamma_bR_c}). \tag{8.55}$$

The Chernoff bound on the Q-function is $Q(\sqrt{2x}) < e^{-x}$. Applying this bound to (8.54) yields

$$P_e \leq (2^k - 1)e^{-\gamma_bR_cd_{\min}} < 2^ke^{-\gamma_bR_cd_{\min}} = e^{-\gamma_bR_cd_{\min}+k\ln 2}. \tag{8.56}$$

Comparing this bound with that of uncoded BPSK modulation,

$$P_b = Q(\sqrt{2\gamma_b}) < e^{-\gamma_b}, \tag{8.57}$$

we get a dB coding gain of approximately

$$G_c = 10\log_{10}[(\gamma_bR_cd_{\min} - k\ln 2)/\gamma_b] = 10\log_{10}[R_cd_{\min} - (k\ln 2)/\gamma_b]. \tag{8.58}$$

Note that the coding gain depends on the code rate, the number of information bits per code-word, the minimum distance of the code, and the channel SNR. In particular, the coding gain decreases as γ_b decreases, and it becomes negative at sufficiently low SNRs. In general the performance of SDD is about 2–3 dB better than HDD [6, Chap. 8.1].

Example 8.6: Find the approximate coding gain of SDD over uncoded modulation for the (24, 12) code with $d_{\min} = 8$ considered in Example 8.5, with $\gamma_b = 10$ dB.
Solution: Setting $\gamma_b = 10$, $R_c = 12/24$, $d_{\min} = 8$, and $k = 12$ in (8.58) yields $G_c = 5$ dB. This significant coding gain is a direct result of the large minimum distance of the code.

8.2.8 Hamming, Golay, and BCH Codes

We now describe some common linear block codes. More details can be found in [6, 7, 8, 1]. The most common type of block code is a Hamming code, which is parameterized by an integer $m \geq 2$. For an $(n,\ k)$ Hamming code, $n = 2^m - 1$ and $k = 2^m - m - 1$, so $n - k = m$ redundant bits are introduced by the code. The minimum distance of all Hamming codes is $d_{\min} = 3$, so $t = 1$ error in the $n = 2^m - 1$ codeword symbols can be corrected. Although Hamming codes are not very powerful, they are perfect codes and thus have probability of error given exactly by the right side of (8.44).

Golay and extended Golay codes are another class of binary linear block codes with good performance. The Golay code is a linear (23, 12) code with $d_{\min} = 7$ and $t = 3$. The extended Golay code is obtained by adding a single parity bit to the Golay code, resulting in a $(24, 12)$ block code with $d_{\min} = 8$ and $t = 3$. The extra parity bit does not change the error correction capability (since t remains the same), but it greatly simplifies implementation because the information bit rate is exactly half the coded bit rate. Thus, both uncoded and coded bit streams can be generated by the same clock, using every other clock sample to generate the uncoded bits. These codes have higher $d_{\min}$ and thus better error correction capabilities than Hamming codes, but at a cost of more complex decoding and a lower code rate $R_c = k/n$. The lower code rate implies that the code either has a lower data rate or requires additional bandwidth.

The most powerful class of binary linear block codes is the Bose–Chadhuri–Hocquenghem (BCH) codes. These are cyclic codes, and at high rates they typically outperform all other binary linear block codes with the same n and k at moderate to high SNRs. This code class provides a large selection of blocklengths, code rates, and error correction capabilities. In particular, the most common BCH codes have $n = 2^m - 1$ for any integer $m \geq 3$.

The P_b for a number of BCH codes under hard decision decoding and BPSK modulation is shown in Figure 8.6. The plot is based on the approximation (8.50), where for BPSK we have

$$p = Q\left(\sqrt{\frac{2E_c}{N_0}}\right) = Q\left(\sqrt{2R_c\gamma_b}\right). \tag{8.59}$$

In the figure, the BCH (127, 36) code actually has a negative coding gain at low SNRs. This is not uncommon for powerful channel codes owing to their reduced energy per symbol, as discussed in Section 8.2.6.

8.3 Nonbinary Block Codes: The Reed Solomon Code

A nonbinary block code has similar properties as the binary code: it has K information symbols mapped into codewords of length N. However, the N codeword symbols of each codeword are chosen from a nonbinary alphabet of size $q > 2$. Thus the codeword symbols can take any value in $\{0, 1,\ \ldots,\ q - 1\}$. Usually $q = 2^k$, so that k bits can be mapped into one symbol.

The most common nonbinary block code is the Reed Solomon (RS) code, which are a class of non-binary BCH codes. RS codes have $N = q - 1 = 2^k - 1$ and $K = 1, 2, \ldots, N - 1$. The value of K dictates the error correction capability of the code. Specifically, an RS code can correct up to $t = .5\lfloor N - K \rfloor$ codeword symbol errors. In nonbinary codes the minimum distance between codewords is defined as the number of codeword

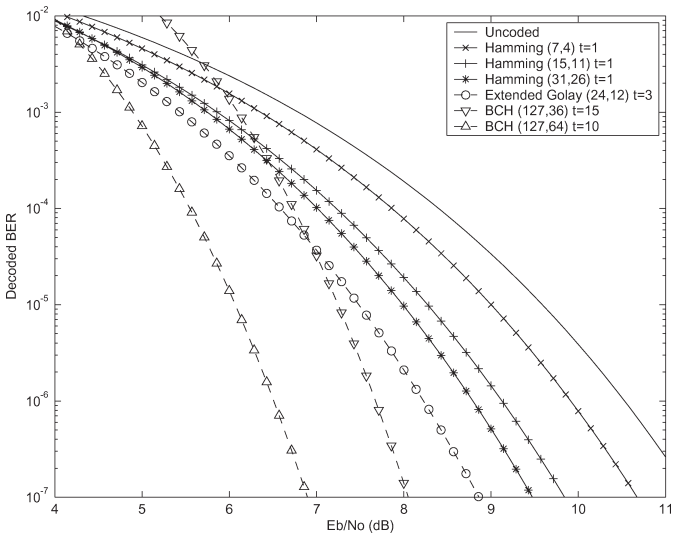


Figure 8.6: P_b for Hamming, Extended Golay, and BCH Codes.

symbols in which the codewords differ. RS codes achieve a minimum distance of $d_{\min} = N - K + 1$, which is the largest possible minimum distance between codewords for any linear code with the same encoder input and output block lengths. Reed Solomon codes are often *shortened* to meet the requirements of a given system [1, Chap. 5.10].

The nonbinary nature of RS codes makes them inferior to binary codes when used with binary modulation. That is because each RS symbol corresponds to $M = 2^k$ coded bits, and any one of these coded bits received in error results in an error for the RS symbol. On the other hand, if all M coded bits are received in error, this also results in a single RS symbol error. Since a burst of M errors has the same effect as a single error in RS codes transmitted using binary modulation, RS codes are effectively optimal binary burst-error-correcting codes. This property has led to their widespread use in mass storage systems such as optimal and magnetic disks to correct bursts of errors associated with media defects. These burst-error-correction benefits also make RS codes commonly used in the outer code of a concatenated coding system, as described in 8.5 below.

Because nonbinary codes – and RS codes in particular – generate symbols corresponding to $M = 2^k$ bits, they are sometimes used with M-ary modulation. In particular, with 2^k-ary modulation each codeword symbol is transmitted over the channel as one of 2^k possible constellation points. If the error probability associated with the modulation (the probability of mistaking the received constellation point for a constellation point other than the transmitted point) is P_M, then the probability of codeword error associated with the non-binary code is upper bounded by

$$P_e \leq \sum_{j=t+1}^{N} \binom{N}{j} P_M^j (1 - P_M)^{N-j}, \tag{8.60}$$

which is similar to the form of (8.44) for the binary code. We can then approximate the probability of information symbol error as

$$P_s \approx \frac{1}{N} \sum_{j=t+1}^{N} j \binom{N}{j} P_M^j (1 - P_M)^{N-j}. \tag{8.61}$$

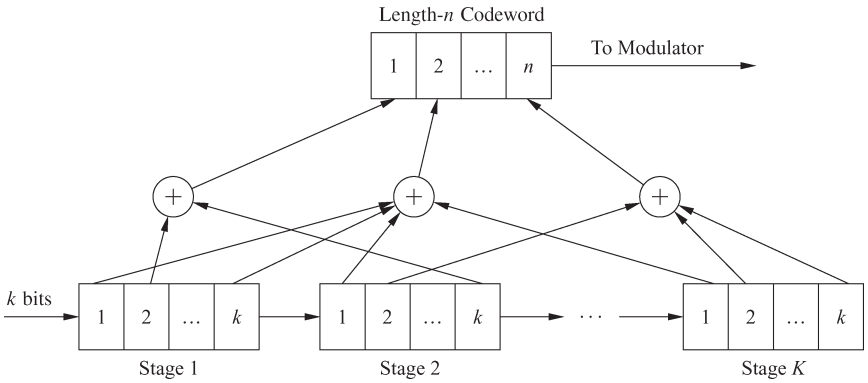


Figure 8.7: Convolutional encoder.

8.4 Convolutional Codes

Convolutional codes were invented by Peter Elias in 1955 [11]. Elias's goal in this seminal paper was to find classes of codes for the binary symmetric channel with error performance approaching that of Shannon's random codes, yet with structure that enabled low-complexity encoding and decoding. It is in fact quite remarkable that convolution codes were invented less than a decade after Shannon's landmark paper, yet they remain widely used in modern communication systems.

A convolutional code generates coded symbols by passing the information bits through a linear finite-state shift register, as shown in Figure 8.7. The shift register consists of K stages with k bits per stage. There are n binary addition operators with inputs taken from all K stages: these operators produce a codeword of length n for each k-bit input sequence. Specifically, the binary input data is shifted into each stage of the shift register k bits at a time, and each of these shifts produces a coded sequence of length n. The rate of the code is $R_c = k/n$. The maximum span of output symbols that can be influenced by a given input bit in a convolutional code is called the *constraint length* of the code. It is clear from Figure 8.7 that a length-n codeword depends on kK input bits–in contrast to a block code, which only depends on k input bits. Thus, the constraint length of the encoder is kK bits or, equivalently, K k-bit bytes. Convolutional codes are said to have memory since the current codeword depends on more input bits (kK) than the number input to the encoder to generate it (k). Note that in general a convolutional encoder may not have the same number of bits per stage.

8.4.1 Code Characterization: Trellis Diagrams

When a length-n codeword is generated by the convolutional encoder of Figure 8.7, this codeword depends both on the k bits input to the first stage of the shift register as well as the *state* of the encoder, defined as the contents in the other $K-1$ stages of the shift register. In order to characterize this convolutional code, we must characterize how the codeword generation depends both on the k input bits and the encoder state, which has $2^{k(K-1)}$ possible values. There are multiple ways to characterize convolutional codes, including a tree diagram, state diagram, and trellis diagram [6, Chap. 8.2]. The tree diagram represents the encoder in the form of a tree, where each branch represents a different encoder state and the corresponding encoder output. A state diagram is a graph showing the different states of the encoder and the possible state transitions and corresponding encoder outputs. A trellis diagram uses the fact that the tree representation repeats itself once the number of stages in the tree exceeds the constraint length of the code. The trellis diagram simplifies the tree representation by merging nodes in the tree corresponding to the same encoder state. In this section we focus on the trellis representation of a convolutional code, since this is the most common characterization. The details of trellis diagram representation are best described by an example.

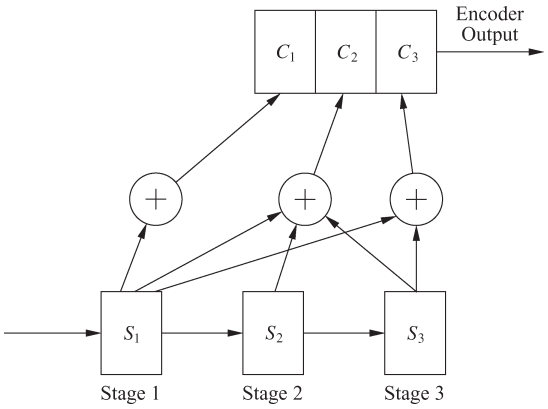


Figure 8.8: Convolutional encoder example ($n = 3,\ k = 1,\ k = 3$).

Consider the convolutional encoder shown in Figure 8.8 with $n = 3, k = 1$, and $K = 3$. In this encoder, one bit at a time is shifted into Stage 1 of the three-stage shift register. At a given time t we denote the bit in Stage i of the shift register as S_i. The three stages of the shift register are used to generate a codeword of length 3, $C_1C_2C_3$; from the figure we see that $C_1 = S_1, C_2 = S_1 + S_2 + S_3$, and $C_3 = S_1 + S_3$. A bit sequence $\mathbf{U}$ shifted into the encoder generates a sequence of coded symbols, which we denote by $\mathbf{C}$. Note that the coded symbols corresponding to C_1 are just the original information bits. As with block codes, when one of the coded symbols in a convolutional code corresponds to the original information bits we say that the code is *systematic*. We define the encoder state as $S = S_2S_3$ (i.e., the contents of the last two stages of the encoder), and there are $2^2 = 4$ possible values for this encoder state. To characterize the encoder, we must show for each input bit and each possible encoder state what the encoder output will be, and we must also show how the new input bit changes the encoder state for the next input bit.

The trellis diagram for this code is shown in Figure 8.9. The solid lines in Figure 8.9 indicate the encoder state transition when a 0-bit is input to Stage 1 of the encoder, and the dashed lines indicate the state transition corresponding to a 1-bit input. For example, starting at state $S = 00$, if a 0-bit is input to Stage 1 then, when the shift register transitions, the new state will remain as $S = 00$ (since the 0 in Stage 1 transitions to Stage 2, and the 0 in Stage 2 transitions to Stage 3, resulting in the new state $S = S_2S_3 = 00$). On the other hand, if a 1-bit is input to Stage 1 then, when the shift register transitions, the new state will become $S = 10$ (since the 1 in Stage 1 transitions to Stage 2, and the 0 in Stage 2 transitions to Stage 3, resulting in the new state $S = S_2S_3 = 10$). The encoder output corresponding to a particular encoder state S and input S_1 is written next to the transition lines in the figure. This output is the encoder output that results from the encoder addition operations on the bits S_1, S_2, and S_3 in each stage of the encoder. For example, if $S = 00$ and $S_1 = 1$ then the encoder output $C_1C_2C_3$ has $C_1 = S_1 = 1, C_2 = S_1 + S_2 + S_3 = 1$, and $C_3 = S_1 + S_3 = 1$. This output 111 is drawn next to the dashed line transitioning from state $S = 00$ to state $S = 10$ in Figure 8.9. Note that the encoder output for $S_1 = 0$ and $S = 00$ is always the all-zero codeword regardless of the addition operations that form the codeword $C_1C_2C_3$, since summing together any number of 0s always yields 0. The portion of the trellis between time t_i and t_{i+1} is called the ith *branch* of the trellis. Figure 8.9 indicates that the initial state at time t_0 is the all-zero state. The trellis achieves *steady state*, defined as the point where all states can be entered from either of two preceding states, at time t_3. After this steady state is reached, the trellis repeats itself in each time interval. Note also that, in steady state, each state transitions to one of two possible new states. In general, trellis structures starting from the all-zero state at time t_0 achieve steady state at time t_K.

For general values of k and K, the trellis diagram will have 2^{K-1} states, where each state has 2^k paths entering each node and 2^k paths leaving each node. Thus, the number of paths through the trellis grows exponentially with

k, K, and the length of the trellis path.

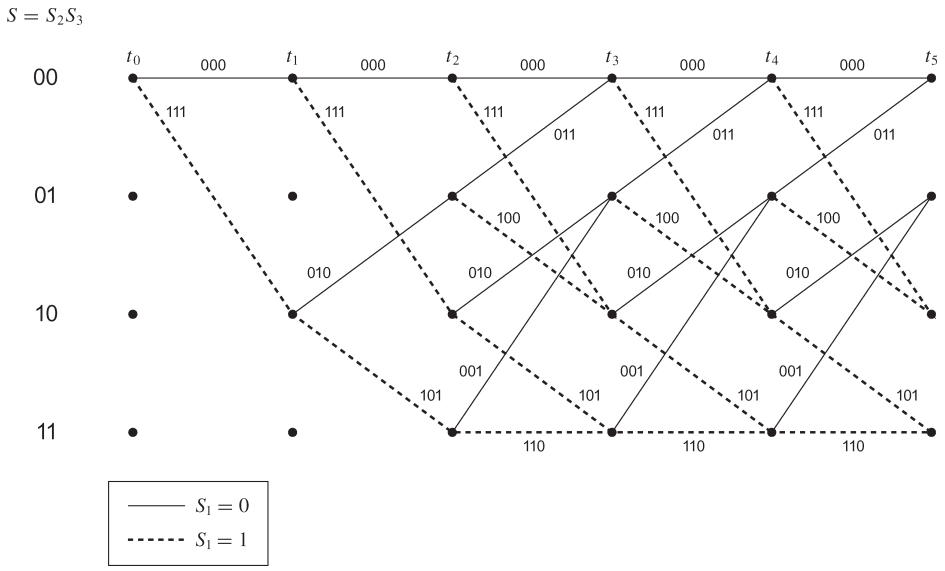


Figure 8.9: Trellis diagram.

Example 8.7: Consider the convolutional code represented by the trellis in Figure 8.7. For an initial state $S = S_2S_3 = 01$, find the state sequence S and the encoder output C for input bit sequence $\mathrm{U} = 011$.
Solution: The first occurrence of $S = 01$ in the trellis is at time t_2. We see at t_2 that if the information bit $S_1 = 0$ then we follow the solid line in the trellis from $S = 01$ at t_2 to $S = 00$ at t_3, and the output corresponding to this path through the trellis is $C = 011$. Now at t_3, starting at $S = 00$, for the information bit $S_1 = 1$ we follow the dashed line in the trellis to $S = 10$ at t_4, and the output corresponding to this path through the trellis is $C = 111$. Finally, at t_4, starting at $S = 10$, for the information bit $S_1 = 1$ we follow the dashed line in the trellis to $S = 11$ at t_5, and the output corresponding to this path through the trellis is $C = 101$.

8.4.2 Maximum Likelihood Decoding

The convolutional code generated by the finite state shift register is basically a finite-state machine. Thus, unlike an $(n,\ k)$ block code – where maximum likelihood detection entails finding the length-n codeword that is closest to the received length-n codeword – maximum likelihood detection of a convolutional code entails finding the most likely sequence of coded symbols $\mathbf{C}$ given the received sequence of coded symbols, which we denote by $\mathbf{R}$. In particular, for a received sequence $\mathbf{R}$, the decoder decides that coded symbol sequence $\mathbf{C}^*$ was transmitted if

$$p(\mathbf{R}|\mathbf{C}^*) \geq p(\mathbf{R}|\mathbf{C}),\ \forall \mathbf{C}. \tag{8.62}$$

Since each possible sequence $\mathbf{C}$ corresponds to one path through the trellis diagram of the code, maximum likelihood decoding corresponds to finding the maximum likelihood path through the trellis diagram. For an AWGN channel, noise affects each coded symbol independently. Thus, for a convolutional code of rate $1/n$, we can express the likelihood (8.62) for a path of length L through the trellis as

$$p(\mathbf{R}|\mathbf{C}) = \prod_{i=0}^{L-1} p(R_i|C_i) = \prod_{i=0}^{L-1} \prod_{j=1}^{n} p(R_{ij}|C_{ij}), \tag{8.63}$$

where C_i is the portion of the code sequence $\mathbf{C}$ corresponding to the ith branch of the trellis, R_i is the portion of the received code sequence $\mathbf{R}$ corresponding to the ith branch of the trellis, C_{ij} is the jth coded symbol corresponding to C_i, and R_{ij} is the jth received coded symbol corresponding to R_i. The log likelihood function is defined as the log of $p(\mathbf{R}|\mathbf{C})$, given as

$$\log p(\mathbf{R}|\mathbf{C}) = \sum_{i=0}^{L-1} \log p(R_i|C_i) = \sum_{i=0}^{L-1} \sum_{j=1}^{n} \log p(R_{ij}|C_{ij}). \tag{8.64}$$

The expression

$$B_i = \sum_{j=1}^{n} \log p(R_{ij}|C_{ij}) \tag{8.65}$$

is called the *branch metric* because it indicates the component of (8.64) associated with the ith branch of the trellis. The sequence or path that maximizes the likelihood function also maximizes the log likelihood function, since the $\log$ is monotonically increasing. However, it is computationally more convenient for the decoder to use the log likelihood function because it involves a summation rather than a product. The log likelihood function associated with a given path through the trellis is also called the *path metric*, which by (8.64) is equal to the sum of branch metrics along each branch of the path. The path through the trellis with the maximum path metric corresponds to the maximum likelihood path.

The decoder can use either hard or soft decisions for the expressions $\log p(R_{ij}|C_{ij})$ in the log likelihood metric. For hard decision decoding, the R_{ij} is decoded as a 1 or a 0. The probability of hard decision decoding error depends on the modulation and is denoted as p. If $\mathbf{R}$ and $\mathbf{C}$ are N symbols long and differ in d places (i.e., their Hamming distance is d), then

$$p(\mathbf{R}|\mathbf{C}) = p^d(1-p)^{N-d} \tag{8.66}$$

and

$$\log p(\mathbf{R}|\mathbf{C}) = -d \log \frac{1-p}{p} + N \log(1-p). \tag{8.67}$$

Since $p < .5$, (8.67) is maximized when d is minimized. So the coded sequence $\mathbf{C}$ with minimum Hamming distance to the received sequence $\mathbf{R}$ corresponds to the maximum likelihood sequence.

In soft decision decoding, the value of the received coded symbols (R_{ij}) are used directly in the decoder, rather than quantizing them to 1 or 0. For example, if the C_{ij} are sent via BPSK over an AWGN channel with a 1 mapped to $\sqrt{E_c}$ and a 0 mapped to $-\sqrt{E_c}$, then

$$R_{ij} = \sqrt{E_c}(2C_{ij} - 1) + n_{ij}, \tag{8.68}$$

where $E_c = kE_b/n$ is the energy per coded symbol and n_{ij} denotes Gaussian noise of mean zero and variance $\sigma^2 = .5N_0$. Thus,

$$p(R_{ij}|C_{ij}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(R_{ij} - \sqrt{E_c}(2C_{ij}-1))^2}{2\sigma^2}\right]. \tag{8.69}$$

Maximizing this likelihood function is equivalent to choosing the C_{ij} that is closest in Euclidean distance to R_{ij}. In determining which sequence $\mathbf{C}$ maximizes the log likelihood function (8.64), any terms that are common to

two different sequences $\mathbf{C}_1$ and $\mathbf{C}_2$ can be neglected, since they contribute the same amount to the summation. Similarly, we can scale all terms in (8.64) without changing the maximizing sequence. Thus, by neglecting scaling factors and terms in (8.69) that are common to any C_{ij}, we can replace $\sum_{j=1}^{n} \log p(R_{ij}|C_{ij})$ in (8.64) with the *equivalent branch metric*

$$\mu_i = \sum_{j=1}^{n} R_{ij}(2C_{ij} - 1) \tag{8.70}$$

and obtain the same maximum likelihood output.

We now illustrate the path metric computation under both hard and soft decisions for the convolutional code of Figure 8.8 with the trellis diagram in Figure 8.9. For simplicity, we will consider only two possible paths through the trellis and compute their corresponding likelihoods for a given received sequence $\mathbf{R}$. Assume we start at time t_0 in the all-zero state. The first path we consider is the all-zero path, corresponding to the all-zero input sequence. The second path we consider starts in state $S = 00$ at time t_0 and transitions to state $S = 10$ at time t_1, then to state $S = 01$ at time t_2, and finally to state $S = 00$ at time t_3, at which point this path merges with the all-zero path. Since the paths and therefore their branch metrics at times $t < t_0$ and $t \geq t_3$ are the same, the maximum likelihood path corresponds to the path whose sum of branch metrics over the branches in which the two paths differ is smaller. From Figure 8.9 we see that the all-zero path through the trellis generates the coded sequence $\mathbf{C}_0 = 000000000$ over the first three branches in the trellis. The second path generates the coded sequence $\mathbf{C}_1 = 111010011$ over the first three branches in the trellis.

Let us first consider hard decision decoding with error probability p. Suppose the received sequence over these three branches is $\mathbf{R} = 100110111$. Note that the Hamming distance between $\mathbf{R}$ and $\mathbf{C}_0$ is 6 while the Hamming distance between $\mathbf{R}$ and $\mathbf{C}_1$ is 4. As discussed previously, the most likely path therefore corresponds to $\mathbf{C}_1$, since it has minimum Hamming distance to $\mathbf{R}$. The path metric for the all-zero path is

$$M_0 = \sum_{i=0}^{2} \sum_{j=1}^{3} \log P(R_{ij}|C_{ij}) = 6 \log p + 3 \log(1-p), \tag{8.71}$$

while the path metric for the other path is

$$M_1 = \sum_{i=0}^{2} \sum_{j=1}^{3} \log P(R_{ij}|C_{ij}) = 4 \log p + 5 \log(1-p). \tag{8.72}$$

Assuming $p \ll 1$, which is generally the case, this yields $M_0 \approx 6 \log p$ and $M_1 \approx 4 \log p$. Since $\log p < 1$, this confirms that the second path has a larger path metric than the first.

Let us now consider soft decision decoding over time t_0 to t_3. Suppose the received sequence (before demodulation) over these three branches, for $E_c = 1$, is $\mathbf{R} = (.8, -.35, -.15, 1.35, 1.22, -.62, .87, 1.08, .91)$. The path metric for the all-zero path is

$$M_0 = \sum_{i=0}^{2} \mu_i = \sum_{i=0}^{2} \sum_{j=1}^{3} R_{ij}(2C_{ij} - 1) = \sum_{i=0}^{2} \sum_{j=1}^{3} -R_{ij} = -5.11, \tag{8.73}$$

and the path metric for the second path is

$$M_1 = \sum_{i=0}^{2} \sum_{j=1}^{3} R_{ij}(2C_{ij} - 1) = 1.91. \tag{8.74}$$

Thus, the second path has a higher path metric than the first. In order to determine if the second path is the maximum likelihood path, we must compare its path metric to that of all other paths through the trellis.

The difficulty with maximum likelihood decoding is that the complexity of computing the log likelihood function (8.64) grows exponentially with the memory of the code, and this computation must be done for every possible path through the trellis. The Viterbi algorithm, discussed in the next section, reduces the complexity of maximum likelihood decoding by taking advantage of the structure of the path metric computation.

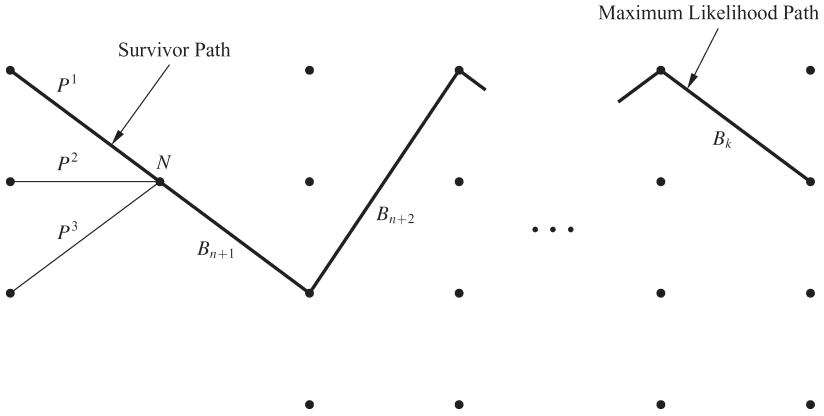


Figure 8.10: Partial path metrics on maximum likelihood path.

8.4.3 The Viterbi Algorithm

The Viterbi algorithm [9], introduced by Viterbi in 1967 [10], reduces the complexity of maximum likelihood decoding by systematically removing paths from consideration that cannot achieve the highest path metric. The basic premise is to look at the partial path metrics associated with all paths *entering* a given node (node N) in the trellis. Since the possible paths through the trellis *leaving* node N are the same for each *entering* path, the complete trellis path with the highest path metric that goes through node N must coincide with the path that has the highest partial path metric up to node N. This is illustrated in Figure 8.10, where path 1, path 2, and path 3 enter node N (at trellis depth n) with partial path metrics $P^l = \sum_{k=0}^{n-1} B_k^l$, $l = 1, 2, 3$, up to this node. Assume P^1 is the largest of these partial path metrics. The complete path with the highest metric has branch metrics $\{B_k\}$ after node N. The maximum likelihood path starting from node N (i.e., the path starting from node N with the largest path metric) has partial path metric $\sum_{k=n}^{\infty} B_k$. The complete path metric for path l, $l = 1, 2, 3$, up to node N and the maximum likelihood path after node N is $P^l + \sum_{k=n}^{\infty} B_k$, $l = 1, 2, 3$, and thus the path with the maximum partial path metric P^l up to node N (path 1 in this example) must correspond to the path with the largest path metric that goes through node N.

The Viterbi algorithm takes advantage of this structure by discarding all paths entering a given node except the path with the largest partial path metric up to that node. The path that is not discarded is called the *survivor path*. Thus, for the example of Figure 8.10, path 1 is the survivor at node N and paths 2 and 3 are discarded from further consideration. Hence, at every stage in the trellis there are 2^{K-1} surviving paths, one for each possible encoder state. A branch for a given stage of the trellis cannot be decoded until all surviving paths at a subsequent trellis stage overlap with that branch; see Figure 8.11, which shows the surviving paths at time t_{k+3}. We see in the figure that all of these surviving paths can be traced back to a *common stem* from time t_k to t_{k+1}. At this point the decoder can output the codeword symbol C_i associated with this branch of the trellis. Note that there is not a fixed decoding delay associated with how far back in the trellis a common stem occurs for a given set of surviving paths, since this delay depends on k, K, and the specific code properties. To avoid a random decoding delay, the

Viterbi algorithm is typically modified so that, at a given stage in the trellis, the most likely branch n stages back is decided upon based on the partial path metrics up to that point. Although this modification does not yield exact maximum likelihood decoding, for n sufficiently large (typically $n \geq 5K$) it is a good approximation.

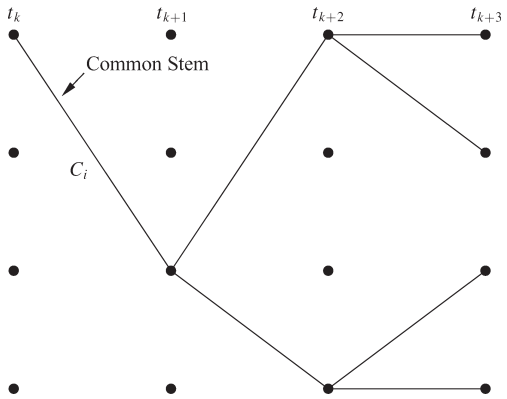


Figure 8.11: Common stem for all survivor paths in the trellis.

The Viterbi algorithm must keep track of $2^{k(K-1)}$ surviving paths and their corresponding metrics. At each stage, in order to determine the surviving path, 2^k metrics must be computed for each node corresponding to the 2^k paths entering each node. Thus, the number of computations in decoding and the memory requirements for the algorithm increase exponentially with k and K. This implies that practical implementations of convolutional codes are restricted to relatively small values of k and K.

8.4.4 Distance Properties

As with block codes, the error correction capability of convolutional codes depends on the distance between codeword sequences. Since convolutional codes are linear, the minimum distance between all codeword sequences can be found by determining the minimum distance from any sequence or, equivalently, determining any trellis path to the all-zero sequence/trellis path. Clearly the trellis path with minimum distance to the all-zero path will diverge and remerge with the all-zero path, so that the two paths coincide except over some number of trellis branches. To find this minimum distance path we must consider all paths that diverge from the all-zero state and then remerge with this state. As an example, in Figure 8.12 we draw all paths in Figure 8.9 between times t_0 and t_5 that diverge and remerge with the all-zero state. Note that path 2 is identical to path 1 – just shifted in time – and thus is not considered as a separate path. Note also that we could look over a longer time interval, but any paths that diverge and remerge over this longer interval would traverse the same branches (shifted in time) as one of these paths plus some additional branches and would therefore have larger path metrics. In particular, we see that path 4 traverses the trellis branches 00-10-01-10-01-00, whereas path 1 traverses the branches 00-10-01-00. Since path 4 traverses the same branches as path 1 on its first, second, and last transition – and since it has additional transitions – its path metric will be no smaller than the metric of path 1. Thus we need not consider a longer time interval to find the minimum distance path. For each path in Figure 8.12 we label the Hamming distance of the codeword on each branch to the all-zero codeword in the corresponding branch of the all-zero path. By summing up the Hamming distances on all branches of each path, we see that path 1 has a Hamming distance of 6 and that paths 3 and 4 have Hamming distances of 8. Recalling that dashed lines indicate 1-bit inputs while solid lines indicate 0-bit inputs, we see that path 1 corresponds to an input bit sequence from t_0 to t_5 of 10000, path 3 corresponds to an input bit sequence of 11000, and path 4 corresponds to an input bit sequence of 10100. Thus, path 1 results in one bit error relative to the all-zero sequence, and paths 3 and 4 result in two bit errors.

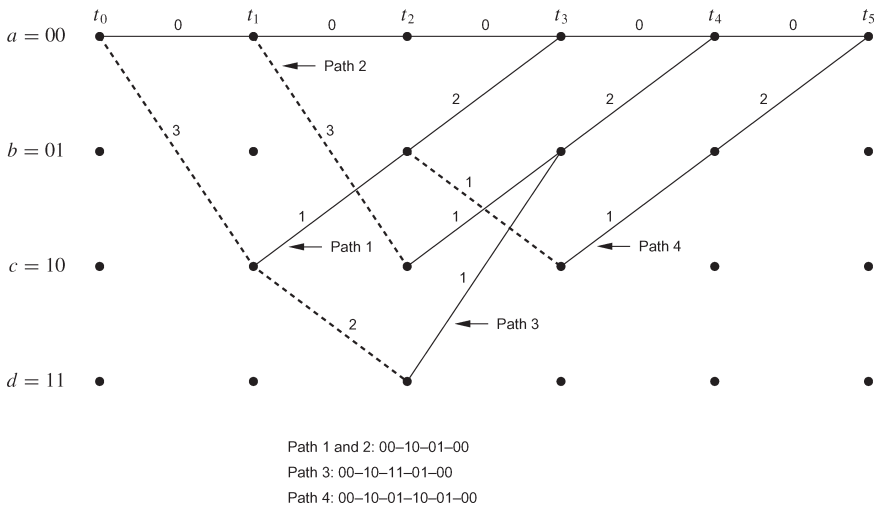


Figure 8.12: Path distances to the all-zero path.

We define the *minimum free distance* d_f of a convolutional code, also called simply the free distance, to be the minimum Hamming distance of all paths through the trellis to the all-zero path, which for this example is 6. The error correction capability of the code is obtained in the same manner as for block codes, with $d_{\min}$ replaced by d_f, so that the code can correct t channels errors with $t = \lfloor .5d_f \rfloor$.

8.4.5 State Diagrams and Transfer Functions

The transfer function of a convolutional code is used to characterize paths that diverge and remerge from the all-zero path, and it is also used to obtain probability of error bounds. The transfer function is obtained from the code's state diagram representing possible transitions from the all-zero state to the all-zero state. The state diagram for the code illustrated in Figure 8.9 is shown in Figure 8.13, with the all-zero state $a = 00$ split into a second node e to facilitate representing paths that begin and end in this state. Transitions between states due to a 0 input bit are represented by solid lines, while transitions due to a 1 input bit are represented by dashed lines. The branches of the state diagram are labeled as either $D^0 = 1, D^1, D^2$, or D^3, where the exponent of D corresponds to the Hamming distance between the codeword (which is shown for each branch transition) and the all-zero codeword in the all-zero path. The self-loop in node a can be ignored because it does not contribute to the distance properties of the code.

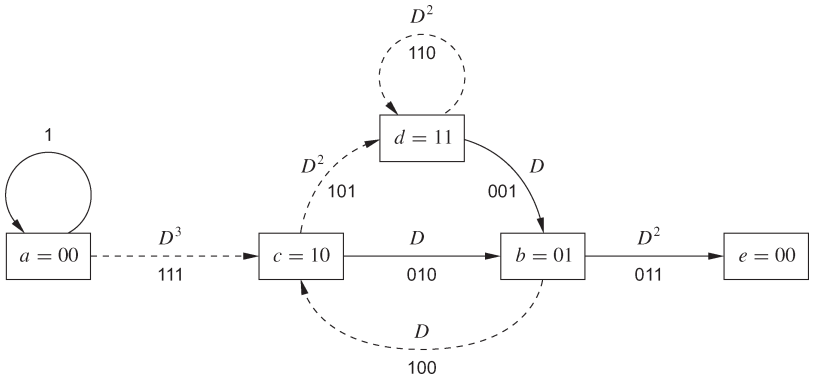


Figure 8.13: State diagram.

The state diagram can be represented by state equations for each state. For the example of Figure 8.13, we obtain state equations corresponding to the four states

$$X_c = D^3 X_a + D X_b, \quad X_b = D X_c + D X_d, \quad X_d = D^2 X_c + D^2 X_d, \quad X_e = D^2 X_b, \tag{8.75}$$

where $X_a, \ldots, X_e$ are dummy variables characterizing the partial paths. The transfer function of the code, describing the paths from state a to state e, is defined as $T(D) = X_e / X_a$. By solving the state equations for the code, which can be done using standard techniques such as Mason's formula, we obtain a transfer function of the form

$$T(D) = \sum_{d=d_f}^{\infty} a_d D^d, \tag{8.76}$$

where a_d is the number of paths with Hamming distance d from the all-zero path. As stated before, the minimum Hamming distance to the all-zero path is d_f, and the transfer function $T(D)$ indicates that there are a_{d_f} paths with this minimum distance. For the example of Figure 8.13, we can solve the state equations given in (8.75) to get the transfer function

$$T(D) = \frac{D^6}{1 - 2D^2} = D^6 + 2D^8 + 4D^{10} + \cdots . \tag{8.77}$$

We see from the transfer function that there is one path with minimum distance $d_f = 6$ and two paths with Hamming distance 8, which is consistent with Figure 8.12. The transfer function is a convenient shorthand for enumerating the number and corresponding Hamming distance of all paths in a particular code that diverge and later remerge with the all-zero path.

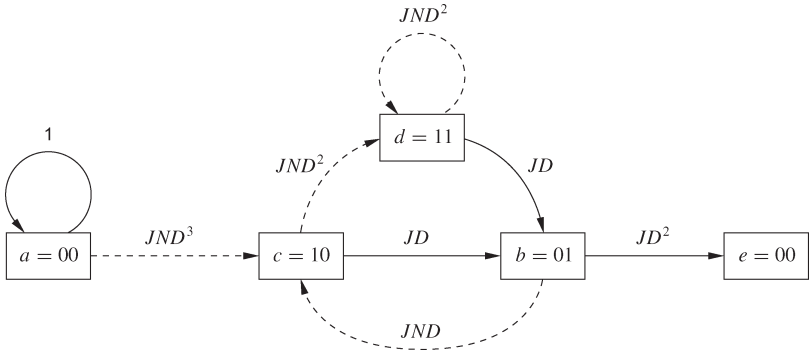


Figure 8.14: Extended state diagram.

Although the transfer function is sufficient to capture the number and Hamming distance of paths in the trellis to the all-zero path, we need a more detailed characterization to compute the bit error probability of the convolutional code. We therefore introduce two additional parameters into the transfer function, N and J, for this additional characterization. The factor N is introduced on all branch transitions associated with a 1 input bit (dashed lines in Figure 8.13). The factor J is introduced to every branch in the state diagram such that the exponent of J in the transfer function equals the number of branches in any given path from node a to node e. The extended state diagram corresponding to the trellis of Figure 8.9 is shown in Figure 8.14.

The extended state diagram can also be represented by state equations. For the example of Figure 8.14, these are given by:

$$X_c = JND^3X_a + JNDX_b, \quad X_b = JDX_c + JDX_d,$$
$$X_d = JND^2X_c + JND^2X_d, \quad X_e = JD^2X_b. \tag{8.78}$$

Similarly to the previous transfer function definition, the transfer function associated with this extended state is defined as $T(D, N, J) = X_e/X_a$, which for this example yields

$$\begin{aligned} T(D, N, J) &= \frac{J^3ND^6}{1 - JND^2(1 + J)} \\ &= J^3ND^6 + J^4N^2D^8 + J^5N^2D^8 + J^5N^3D^{10} + \cdots . \end{aligned} \tag{8.79}$$

The factor J is most important when we are interested in transmitting finite-length sequences; for infinite-length sequences we typically set $J = 1$ to obtain the transfer function for the extended state:

$$T(D, N) = T(D, N, J = 1). \tag{8.80}$$

The transfer function for the extended state tells us more information about the diverging and remerging paths than just their Hamming distance; namely, the minimum distance path with Hamming distance 6 is of length 3 and results in a single bit error (exponent of N is unity), one path of Hamming distance 8 is of length 4 and results in two bit errors, and the other path of Hamming distance 8 is of length 5 and results in two bit errors, consistent with Figure 8.12. The extended transfer function is a convenient shorthand for representing the Hamming distance, length, and number of bit errors that correspond to each diverging and remerging path of a code from the all-zero path. In the next section we show that this representation is useful in characterizing the probability of error for convolutional codes.

8.4.6 Error Probability for Convolutional Codes

Since convolutional codes are linear codes, the probability of error can be obtained by first assuming that the all-zero sequence is transmitted and then determining the probability that the decoder decides in favor of a different sequence. We will consider error probability for both hard decision and soft decision decoding; soft decisions are much more common in wireless systems owing to their superior performance.

First consider soft decision decoding. We are interested in the probability that the all-zero sequence is sent but a different sequence is decoded. If the coded symbols output from the convolutional encoder are sent over an AWGN channel using coherent BPSK modulation with energy $E_c = R_cE_b$, then it can be shown [6] that, if the all-zero sequence is transmitted, the probability of mistaking this sequence with a sequence Hamming distance d away is

$$P_2(d) = Q\left(\sqrt{\frac{2E_c}{N_0}d}\right) = Q(\sqrt{2\gamma_bR_cd}). \tag{8.81}$$

We call this probability the *pairwise error probability*, since it is the error probability associated with a pairwise comparison of two paths that differ in d bits. The transfer function enumerates all paths that diverge and remerge with the all zero path, so by the union bound we can upper bound the probability of mistaking the all-zero path for another path through the trellis as

$$P_e \leq \sum_{d=d_f}^{\infty} a_dQ(\sqrt{2\gamma_bR_cd}), \tag{8.82}$$

where a_d denotes the number of paths of distance d from the all-zero path. This bound can be expressed in terms of the transfer function itself if we use the Chernoff upper bound for the Q-function, which yields

$$Q(\sqrt{2\gamma_b R_c d}) \leq e^{-\gamma_b R_c d}.$$

Using this in (8.82) we obtain the upper bound

$$P_e < T(D)|_{D=e^{-\gamma_b R_c}}. \tag{8.83}$$

This upper bound tells us the probability of mistaking one sequence for another, but it does not yield the more fundamental probability of bit error. We know that the exponent in the factor N of $T(D, N)$ indicates the number of information bit errors associated with selecting an incorrect path through the trellis. Specifically, we can express $T(D, N)$ as

$$T(D, N) = \sum_{d=d_f}^{\infty} a_d D^d N^{f(d)}, \tag{8.84}$$

where $f(d)$ denotes the number of bit errors associated with a path of distance d from the all-zero path. Then we can upper bound the bit error probability for $k = 1$ as

$$P_b \leq \sum_{d=d_f}^{\infty} a_d f(d) Q(\sqrt{2\gamma_b R_c d}) \tag{8.85}$$

(see [6, Chap. 8.2]), which differs from (8.82) only in the weighting factor $f(d)$ corresponding to the number of bit errors in each incorrect path. If the Q-function is upper bounded using the Chernoff bound as before, we get the upper bound

$$P_b < \left.\frac{dT(D, N)}{dN}\right|_{N=1, D=e^{-\gamma_b R_c}}. \tag{8.86}$$

If $k > 1$ then we divide (8.85) or (8.86) by k to obtain P_b.

All of these bounds assume coherent BPSK transmission (or coherent QPSK, which is equivalent to two independent BPSK transmissions). For other modulations, the pairwise error probability $P_2(d)$ must be recomputed based on the probability of error associated with the given modulation.

Let us now consider hard decision decoding. The probability of selecting an incorrect path at distance d from the all-zero path, for d odd, is given by

$$P_2(d) = \sum_{k=.5(d+1)}^{d} \binom{d}{k} p^k (1-p)^{d-k}, \tag{8.87}$$

where p is the probability of error on the channel. This follows because the incorrect path will be selected only if the decoded path is closer to the incorrect path than to the all-zero path – that is, the decoder makes at least $.5(d + 1)$ errors. If d is even then the incorrect path is selected when the decoder makes more than $.5d$ errors, and the decoder makes a choice at random if the number of errors is exactly $.5d$. We can bound the pairwise error probability as

$$P_2(d) < [4p(1-p)]^{d/2}. \tag{8.88}$$

Following the same approach as in soft decision decoding, we then obtain the error probability bound as

$$P_e < \sum_{d=d_f}^{\infty} a_d [4p(1-p)]^{d/2} < T(D)|_{D=\sqrt{4p(1-p)}}, \tag{8.89}$$

and

$$P_b < \sum_{d=d_f}^{\infty} a_d f(d) P_2(d) = \left. \frac{dT(D,N)}{dN} \right|_{N=1,\ D=\sqrt{4p(1-p)}}. \tag{8.90}$$

8.5 Concatenated Codes

A concatenated code uses two levels of coding: an inner code and an outer code, as shown in Figure 8.15. Concatenated codes were invented by G. David Forney Jr. in 1965 [12] to reduce complexity in the decoding of long codes. Prior to Forney's invention, while strong error-correction codes with error probability that decreased exponentially with blocklength were known, the decoding complexity also increased exponentially with blocklength, making decoding of these codes impractical. Forney eliminated this complexity barrier by creating long codes with a structure that could reduce decoding complexity to be polynomial in blocklength rather than exponential.

The basic premise of a concatenated code is to create a long error-correction code using lower-complexity inner and outer code segments. Decoding is done sequentially on each segment, as indicated in the figure; first the inner code is decoded, and then the outer code is decoded. The suboptimal sequential decoding of the inner and outer code replaces the optimal decoder, which performs maximum likelihood (ML) decoding of the concatenated code through a joint decoding of the inner and outer codes. Although suboptimal, with properly designed inner and outer codes sequential decoding can obtain the desired performance for many wireless systems with much lower complexity than is possible with the optimal ML decoder. Another approach to decoding a concatenated code is iterative decoding, which iterates between decoding the inner and outer codes [13]. The first step of iterative decoding is the same as sequential decoding: the inner code is decoded, followed by the outer code. Next, the output of the outer decoder (which may use soft or hard decoding) is used for another round of sequential decoding, and this process can continue for any number of iterations. Since errors in both the inner and outer codes can be corrected with each iteration, iterative decoding can achieve performance close to that of the optimal joint decoder.

There are many types of inner codes used in concatenated codes, depending on the channel characteristics, the complexity constraints, and the desired performance. Concatenated codes designed for wireless channels typically use a convolutional code for the inner code with Viterbi decoding, due to the low decoding complexity and good error-correction properties. Viterbi decoding of convolutional codes tends to have errors that occur in bursts, and burst errors can also be introduced for any inner code when the wireless channel experiences a deep fade. Concatenated codes designed for wireless channels typically use an outer Reed Solomon (RS) code to correct for such error bursts. In fact the large majority of concatenated codes studied and implemented to date use the Reed Solomon code as the outer code, including the first concatenated codes proposed in Forney's 1965 paper. The combination of a convolutional inner code and RS outer code has been widely used for deep space and satellite communications, starting with the 1977 Nasa Voyager mission. As more powerful codes were developed, the inner codes in concatenated codes were replaced. For example, the second generation digital video broadcast standard DVB-S2 replaced the convolutional inner code with an LDPC code, but retained the RS outer code to correct for error bursts. In terrestrial wireless systems where fading can introduce large error bursts, concatenated codes often have the inner and outer codes separated by an interleaver, as shown in the figure. Interleaver design for different coding techniques is described in Section 8.10.

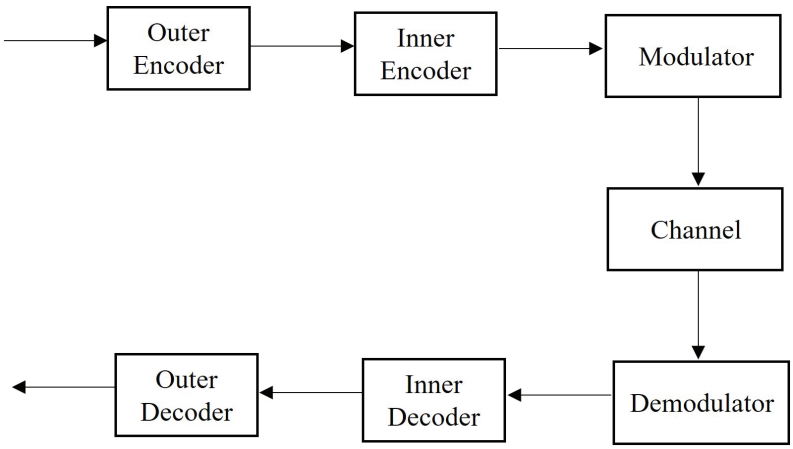


Figure 8.15: Concatenated coding.

The basic premise of concatenated coding can be extended to more than two codes. This extension is referred to as multi-level concatenated coding, where $M > 2$ codes are concatenated together through sequential encoding and decoding for each of the M codes. As in the case of two-level concatenated coding, sequential decoding of more than two concatenated codes is suboptimal, and hence iterative decoding is often used which can achieve close to optimal performance. Multilevel code design entails determining the number and type of codes to use, their order, the decoding technique. These codes are not in widespread use today as they have been largely replaced with the more powerful techniques of turbo and LDPC coding. In fact turbo codes, like concatenated codes, utilize segment codes and iterative decoding, and can do so in parallel or sequentially, as described in the next section.

8.6 Turbo Codes

Turbo codes, published in a landmark 1993 paper by Berrou, Glavieux, and Thitimajshima ([14]; see also [15]), are powerful codes that come within a fraction of a decibel of the Shannon capacity limit on AWGN channels. Turbo codes have been studied extensively since their introduction given their exceptional performance. Turbo codes fall within the general category of codes on graphs, which were introduced by Gallager in the early sixties [16]; at the time, however, these coding techniques were thought impractical and were generally not pursued by researchers in the field. The publication of the 1993 turbo code paper, based on a 1991 patent by Berrou, provided more than enough motivation to revisit the work of Gallager and others on iterative, graph-based decoding techniques [17, 18] and to apply such ideas to the design and analysis of turbo codes [19, 20, 21]

As first described by Berrou et al., turbo codes consist of two key components: parallel concatenated encoding and iterative, "turbo" decoding. A typical parallel concatenated encoder is shown in Figure 8.16. It consists of two parallel convolutional encoders separated by an interleaver, with the input to the channel being the data bits m along with the parity bits X_1 and X_2 output from each of the encoders in response to input m. Channel codes where the information bits are transmitted as part of the codeword is called a *systematic code*, which is the case for the parallel concantenated code shown in the figure. The key to parallel concatenated encoding lies in the recursive nature of the encoders and the impact of the interleaver on the information stream. Interleavers also play a significant role in the reduction of error floors, which are commonly exhibited in turbo codes.

Iterative or "turbo" decoding exploits the component-code substructure of the turbo encoder by associating a component decoder with each of the component encoders. More specifically, each decoder performs soft input–soft output decoding, as shown in Figure 8.17 for the example encoder of Figure 8.16. In this figure decoder 1 generates

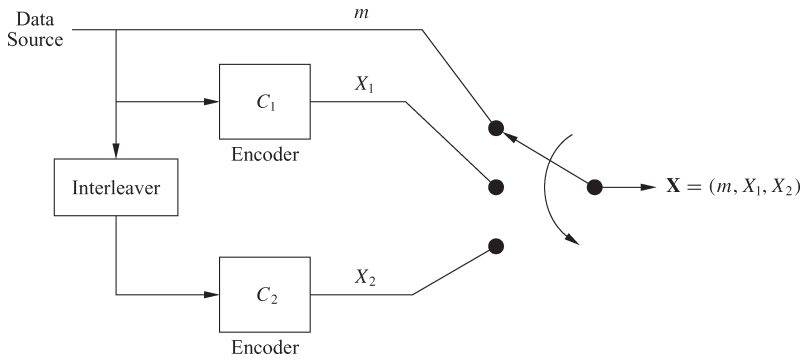


Figure 8.16: Parallel concatenated (turbo) encoder.

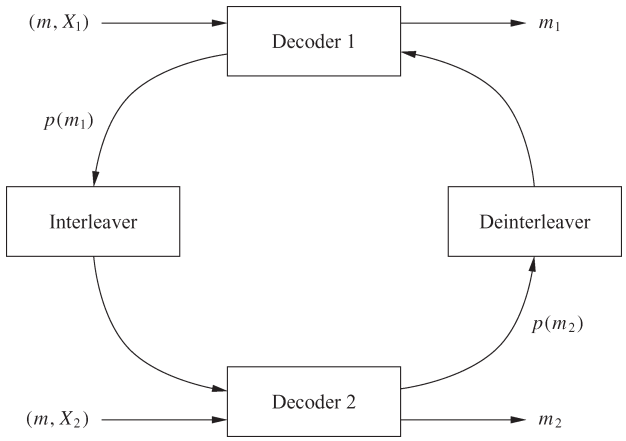


Figure 8.17: Turbo decoder.

a soft decision in the form of a probability measure $p(m_1)$ (or its log value) on the transmitted information bits based on the received codeword (m, X_1). The probability measure is typically generated using the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm, which is based on the maximum a posteriori (MAP) probability [22], or using a soft output Viterbi algorithm (SOVA), which attaches a reliability indicator to the VA hard decision outputs [1, Chap. 12.5]. This probability information is passed to decoder 2, which generates its own probability measure $p(m_2)$ from its received codeword (m, X_2) and the probability measure $p(m_1)$. This reliability information is input to decoder 1, which revises its measure $p(m_1)$ based on this information and the original received codeword. Decoder 1 sends the new reliability information to decoder 2, which revises its measure using this new information. Turbo decoding proceeds in an iterative manner, with the two component decoders alternately updating their probability measures. At the end of the iterative process, the output of the second decoder determines the iterative decoder output.

With a large number of iterations the two decoders generally converge towards the same probability measure $p(m_1) = p(m_2)$, the same soft decision $m_1 = m_2$, and the same probability of error. However, each iteration entails additional delay and energy consumption, while the performance gains associated with each iteration diminish. Hence, turbo decoders generally operate based on a fixed number of iterations or a stopping criteria with some maximum bound on iterations. In the case of a fixed number of iterations, the number is determined based on the tradeoff between delay, energy consumption, and bit error rate performance. Extensive work on iterative turbo decoding performance for different component encoders and decoders has shown that, for AWGN channels, more than 8-12 iterations leads to negligible performance gain. However, iterative decoders based on a fixed number of

iterations entail unnecessary delay and energy consumption when the decoding converges before the fixed number of iterations is reached. To address this issue, different stopping criteria for turbo decoders have been proposed, whereby the iterative process terminates when the stopping criteria is achieved, up to some maximum number of iterations. Many such stopping criteria have been proposed based either on the soft or hard outputs of the component decoders. Examples of early stopping criteria include requiring a minimum cross-entropy between the soft or hard outputs of the two decoders, or requiring a minimum variance of $p(m_i)$ or its log-likelihood for one of the component decoders [20]. A survey on more recent stopping criteria can be found in [23]

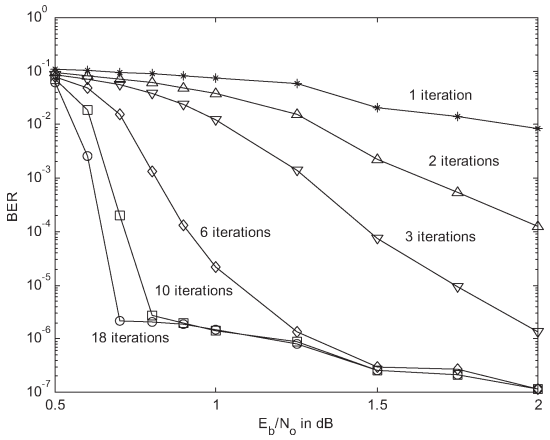


Figure 8.18: Turbo code performance: rate $1/2$, $K = 5$ component codes with interleaver depth 2^{16} from [24].

The simulated performance of turbo codes over multiple iterations of the decoder, as determined by Valenti in [24], is shown in Figure 8.18 for a code composed of two convolutional codes of rate $1/2$ with constraint length $K = 5$ separated by an interleaver of depth $d = 2^{16} = 65536$. The decoder converges after approximately eighteen iterations. This curve indicates several important aspects of turbo codes. First, note their exceptional performance: bit error probability of 10^{-6} at an E_b/N_0 of less than 1 dB. In fact, the original turbo code proposed in [14] performed within .5 dB of the Shannon capacity limit at $P_b = 10^{-5}$. The intuitive explanation for the amazing performance of turbo codes is that the code complexity introduced by the encoding structure is similar to the codes that achieve Shannon capacity. The iterative procedure of the turbo decoder allows these codes to be decoded without excessive complexity. However, note that the turbo code exhibits an error floor: in Figure 8.18 this floor occurs at 10^{-6}. This floor is problematic for systems that require extremely low bit error rates. Several mechanisms have been investigated to lower the error floor, including bit interleaving and increasing the constraint length of the component codes.

An alternative to parallel concatenated turbo encoding and decoding is to encode and decode the component codes in a serial manner [25]. In this coding technique, one component code serves as an outer code, and then the output of this first encoder is interleaved and passed to a second encoder. The output of the second encoder comprises the coded bits. Decoding and deinterleaving is iterated between the inner and outer codes with soft information passed between them. There has been much work comparing turbo codes based on serial versus parallel encoding and decoding (see e.g. [25, 26, 27]). Whereas both codes perform very well under similar delay and complexity conditions, in some cases turbo codes based on serial encoding and decoding perform better at low bit error rates and also can exhibit a lower error floor.

8.7 Low-Density Parity-Check Codes

Low-density parity-check (LDPC) codes were originally invented by Gallager [18]. However, these codes were largely ignored until the introduction of turbo codes, which rekindled some of the same ideas. Subsequent to the landmark paper [14] on turbo codes in 1993, LDPC codes were announced by Mackay and Neal [28] and Wiberg [29]. Shortly thereafter it was recognized that these new code designs were actually reinventions of Gallager's original ideas, and subsequently much work has been devoted to finding the capacity limits, encoder and decoder designs, and practical implementation of LDPC codes for different channels.

LDPC codes are a special class of systematic linear block codes, discussed in Section 8.2.3, with the property that the parity check matrix of the code is sparse. Recall that the codewords in systematic linear block codes consist of information bits and parity bits, corresponding to a special structure of the generator matrix $\mathbf{G} = [\mathbf{I}_k|\mathbf{P}]$. The code is defined by its generator matrix or the corresponding parity check matrix given by $\mathbf{H} = [\mathbf{P}^T|\mathbf{I}_{n-k}]$. The parity check matrix satisfies the property that $\mathbf{CH^T} = \mathbf{0}$ for any valid codeword $\mathbf{C}$. Hence, it is easily determined if a received codeword has been corrupted by simply multiplying it with the transpose of the parity check matrix. The parity check matrix in an (n, k) systematic linear block code is *sparse* if the number of 1s in each column and the number of 1s in each row are both much smaller than the codeword length n. Since the fraction of nonzero entries in $\mathbf{H}$ is small, the parity-check matrix for the code has a low density – hence the name low-density parity-check codes. An LDPC code is *regular* if both the number of 1s in each column is the same and the number of 1st in each row is the same, i.e. the parity-check matrix $\mathbf{H}$ has d_v 1s in each column and d_c ls in each row, where the sparsity condition implies $d_v << n$ and $d_c << n$. If an LDPC code is not regular, it is called irregular.

Another way to define an LDPC code is through the *Tanner Graph* representation of its parity check matrix $\mathbf{H}$ [30]. The Tanner graph for the parity check matrix in an (n, k) systematic linear block code consists of two types of nodes, n variable (or bit) nodes which in the graph are denoted as $v_i, i = 1, \ldots, n$ and represented as circles, and $n - k$ check nodes, which in the graph are denoted as $c_i, i = 1, \ldots, n - k$ and represented as squares. The edges of the Tanner graph indicate the non-zero entries of the parity check matrix $\mathbf{H}$. In particular, variable node v_j is connected to check node c_i if $h_{ij} = 1$ in the parity check matrix $\mathbf{H}$, which indicates that variable node v_j is used to compute parity check node c_i. The nodes of a Tanner graph and edges between them for a $(5, 2)$ code are shown in Figure 8.19. The Tanner graph for a regular LDPC codes has the property that each variable node has the same number of incident edges and each check node has the same number of incident edges. Irregular LDPC codes have different numbers of incident edges for the variable and/or check nodes. The sparsity condition of LDPC parity check matrices imply that there are a small number of edges in the code's Tanner graph, on the order of the code length n, and far fewer than the $n(n - k)$ edges that would occur if the graph were fully connected.

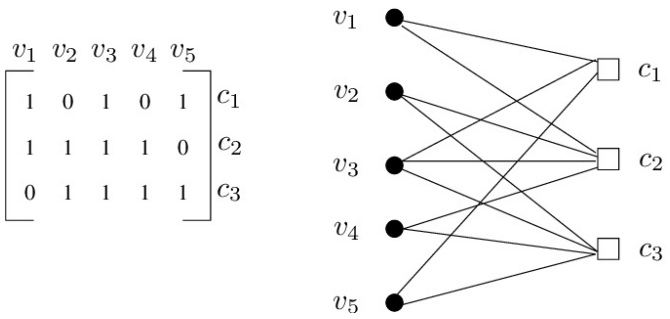


Figure 8.19: Tanner Graph Representation of the Parity Check Matrix for a $(5, 2)$ Systematic Linear Block Code

A set of edges between nodes in a Tanner graph form a loop or cycle when traversing these edges leads back to the original node. For example, in Figure 8.19 traversing edges v_2 to c_2 to v_4 to c_3 to v_2 forms a cycle. The *girt* of a Tanner graph is its shortest cycle. Since LDPC codes are low density, there are relatively few edges in their

Tanner graphs, and hence an (n, k) LDPC code will generally have a much large girth than an (n, k) which is not low-density. This low-density characteristic can be exploited to reduce the complexity for LDPC decoding, as was originally proposed by Gallager in his landmark paper [18].

A common method to decode an LDPC code is called the message passing algorithm (MPA) [31], which is an iterative algorithm based on the code's Tanner graph. In the first half of the MPA iteration, each v node in the Tanner graph computes the probability that the received bit is a 1 or 0 based on the node's input messages, and passes this probability to the c nodes it is connected to via an edge. In the next half iteration each c node determines the probability that its parity check equation is satisfied and passes this information to the v nodes it is connected to via an edge. These iterations repeat for a fixed total number of iterations or until some stopping criteria is met. The MPA assumes that the messages passed are statistically independent, which is only true when the Tanner graph has no cycles. The larger the girth of the Tanner graph, the less degradation in the performance of the MPA decoder relative to an optimal decoder that is caused by the dependence associated with cycles in the graph.

Provided that the codeword length is long, LDPC codes achieve performance close to the Shannon limit and in some cases surpass the performance of turbo codes [32]. The fundamental practical difference between turbo codes and LDPC codes is that turbo codes tend to have low encoding complexity (linear in blocklength) but high decoding complexity (due to their iterative nature and message passing). In contrast, LDPC codes tend to have relatively high encoding complexity but low decoding complexity. In particular, like turbo codes, LDPC decoding uses iterative techniques that are related to Pearl's belief propagation, which is commonly used by the artificial intelligence community (see [33]). However, the belief propagation corresponding to LDPC decoding may be simpler than for turbo decoding [33, 34]. In addition, this belief propagation decoding is parallelizable and can be closely approximated with decoders of very low complexity [31]. Finally, the decoding algorithm for LDPC codes can determine when a correct codeword has been detected, which is not necessarily the case for turbo codes. For all of these reasons, LDPC codes are widely used in practice today.

Additional work in the area of LDPC codes includes finding capacity limits for these codes [31], determining effective code designs [35] and efficient encoding and decoding algorithms [31, 36], and expanding the code designs to include nonregular [32] and nonbinary LDPC codes [37] as well as coded modulation [39]. The textbooks [2, 3, 4] provide detailed treatments of LDPC codes, including code design, decoding algorithm tradeoffs, and LDPC code performance analysis.

8.8 Polar Codes

Polar codes, invented by Arikan [38], are block codes that transform a block of transmissions over a physical discrete memoryless channel W to create virtual channels that, as the blocklength grows, are either highly reliable (asymptotically small error probability) or highly unreliable (probably of error approaching one). The virtual channels are thus polarized in terms of their reliability. The block codes are designed such that the information bits are sent only over the highly reliable virtual channels, while dummy, or *frozen*, bits are sent over the unreliable channels and subsequently discarded by the decoder. Polar codes are the only explicit code construction to achieve capacity of a binary discrete memoryless channel in the limit of infinite blocklength, in contrast to the random capacity-achieving coding scheme developed by Shannon.

The design of polar codes entails creation of virtual channels through the process of channel combining and splitting, as shown in Fig. 8.20. The channel combining transforms N transmissions over the physical channel W with transition probability $W(y|x)$ and capacity $C(W)$ into transmission over a vector channel W^N of dimension N. The channel splitting then maps this vector channel into transmissions over N virtual channels $W^N = W_1..W_N$. Of these virtual channels, $M = NC(W)$ are reliable so M information bits are sent over the M reliable channels, while $N - M$ frozen bits are sent over the unreliable virtual channels.

The virtual channels become polarized through the design of the combiner, which is recursive. The recursive

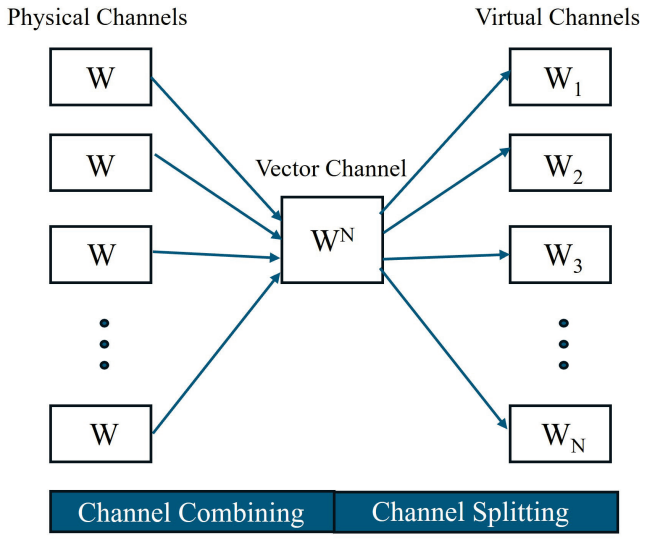


Figure 8.20: Polar Code Channel Combining and Splitting

design increases the reliability of the subset $M = NC(W) < N$ of the N virtual channels, while reducing the reliability of the other $N(1 - C(W))$ channels. As N approaches infinity the error probability of this subset of channels approaches zero, and the error probability of the other channels approaches one. In this limit information bits are only sent over the $NC(W)$ reliable channels via a codeword of length N, leading to a data rate of the channel capacity $C(W)$.

The recursive combining process leading to channel polarization is shown in Fig. 8.21. The physical channel W is considered the first level, so we set $W = W_1$. For channel W_2 we have $x_1 = u_1 \oplus u_2$, $x_2 = u_2$, and transition probabilities $W_2(y_1, y_2|u_1, u_2) = W(y_1|u_1 \oplus u_2)W(y_2|u_2)$. Since the operation mapping (u_1, u_2) to (x_1, x_2) is invertible, mutual information is preserved, i.e. $I(W_2(u_1, u_2|y_1, y_2) = I(W(y_1|u_1 \oplus u_2)) + I(W(y_2|u_2)) = 2I(W)$. Moreover, it is easily shown that $I(W(y_1|u_1 \oplus u_2)) \leq I(W(y_2|u_2))$ and hence we conclude that

$$I(W(y_1|u_1 \oplus u_2)) \leq I(W) \leq I(W(y_2|u_2)). \tag{8.91}$$

Thus we see that the process of channel polarization has already begun with this first iteration, as the capacity for one of the virtual channels is less than that of the other virtual channel.

The next iteration in the recursive combining shown in Fig. 8.21 creates the channel W_4 with $x^4 = u^4 G_4$ for

$$G_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{8.92}$$

and transition probabilities $W_4(y_1, y_2, y_3, y_4|u_1, u_2, u_3, u_4) = W_2(y_1, y_2|u_1 \oplus u_2, u_3 \oplus u_4)W_2(y_3, y_4|u_2, u_4)$. For any given physical channel W, an approximate fraction $I(W)$ of the four virtual channels associated with W_4 will have mutual information less than or equal to $I(W)$ while others will have mutual information greater than or equal to $I(W)$.

The recursion can now be defined for any N that is a power of two, and is shown in Fig. 8.21. In this general case the virtual channel W_N is created from two virtual channels $W_{N/2}$ whose inputs are based on the

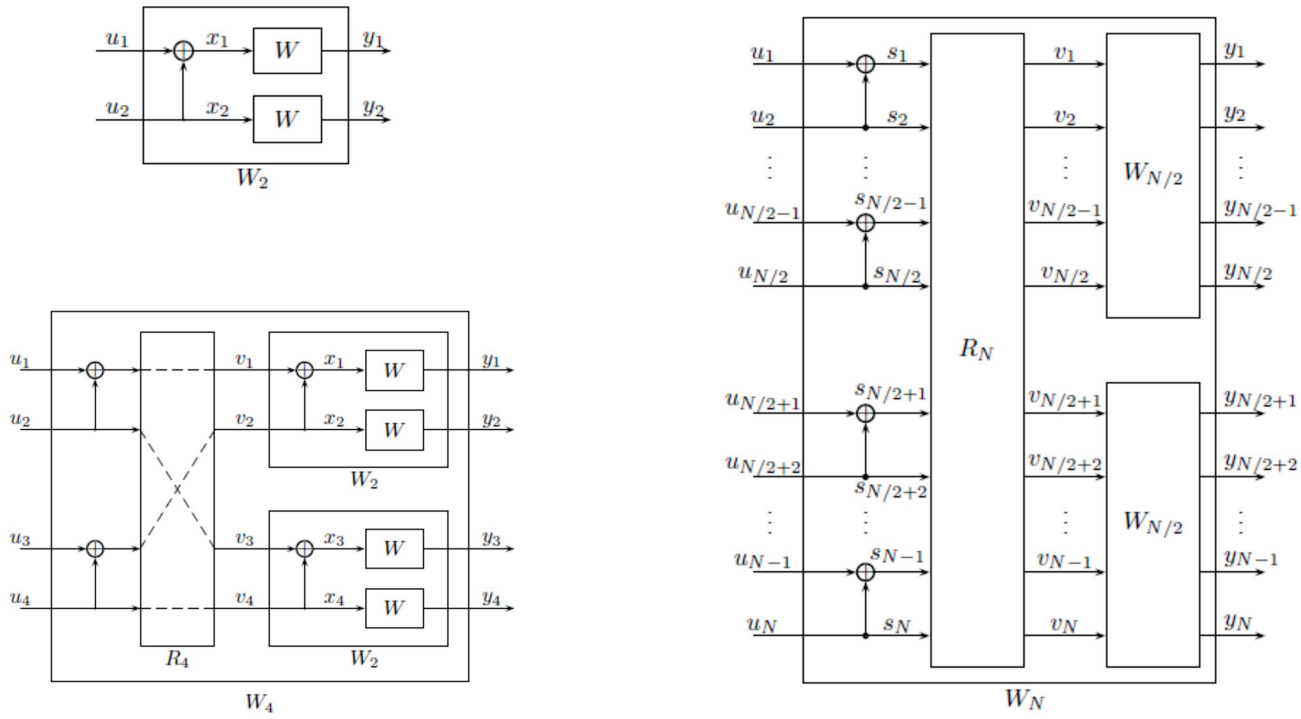


Figure 8.21: Polar Code Channel Combining via Recursion (Fig. 1-3 of [38]

mapping $x^N = u^N G_N$ for the generator matrix $G_N = B_N F$, where B_N is a particular type of permutation matrix called the bit-reversal matrix and F is a structured combining matrix [38]. The transition probability of the virtual channel W_N is related to N transmissions over the physical channel W by $W_N(y^N|u^N) = W^N(y^N|u^N G_N)$. The polarization of the N channels associated with virtual channel W_N is derived in [38] for any discrete memoryless channel W as follows: For the N virtual channels in W_N and for any $\delta \in (0,1)$, as N goes to infinity, a fraction $I(W)$ of the N channels have mutual information within $(1-\delta)$ of $C(W)$ while a fraction $(1-I(W))$ have mutual information between zero and δ. Optimizing the input distribution such that $I(W) = C(W)$ thus achieves channel capacity on the virtual channel W_N in the limit of infinite blocklength.

The creation of virtual channels through channel combining is undone at the receiver through channel splitting. Specifically, the channel output y^N is split to create N binary input, multiple output channels where the ith channel W_N^i has input u_i, output (y^N, u^{i-1}), and transition probability

$$W_N^i(y^N, u^{i-1}|u_i) = \sum_{u_{i+1}^N} \frac{1}{2^N - 1} W_N^i(y^N|u^N). \tag{8.93}$$

Successive decoding is used to decode the split channels, whereby the input u_i is estimated from y^N and the previously-decoded inputs u^{i-1}. This successive decoding proceeds until all of the bits associated with the reliable channels W_1 through $W_{NC(W)}$ are decoded correctly. However, if there is an error in decoding any of these information bits, that error is propagated to all subsequent decoders. Successive decoding in polar codes is optimal when the probability of error for each of the virtual channels is zero but that is only the case in the limit of infinite blocklength. For code blocklengths typically used in practice, which are limited in size due to latency and complexity, successive decoding can have a high error probability due to error propagation. Thus polar codes with successive decoding generally have higher error probability than LDPC and turbo codes. Polar decoding that uses

successive list decoding has better performance than successive decoding, but the list decoding adds significant complexity that limits it practicality. Hence polar codes have generally been used in low data rate applications such as control channels [40, **?**].

8.9 Coded Modulation

Although Shannon proved the capacity theorem for AWGN channels in the late 1940s, it wasn't until the 1990s that rates approaching the Shannon limit were attained – primarily for AWGN channels with binary modulation using turbo codes. Shannon's theorem predicted the possibility of reducing both energy and bandwidth simultaneously through coding. However, as described in Section 8.1, traditional error correction coding schemes (e.g., block, convolutional, and turbo codes) provide coding gain at the expense of increased bandwidth or reduced data rate.

The spectrally efficient coding breakthrough came when Ungerboeck [42] introduced a coded modulation technique to jointly optimize both channel coding and modulation. This joint optimization results in significant coding gains without bandwidth expansion. Ungerboeck's trellis coded modulation, which uses multilevel/phase signal modulation and simple convolutional coding with mapping by set partitioning, compares favorably to later developments in coded modulation (coset and lattice codes) as well as to more complex trellis codes [**?**]. We now outline the general principles of this coding technique. Comprehensive treatments of trellis, lattice, and coset codes can be found in [42, 43, 44, **?**].

The basic scheme for trellis and lattice coding – or, more generally, for any type of coset coding – is depicted in Figure 8.22.

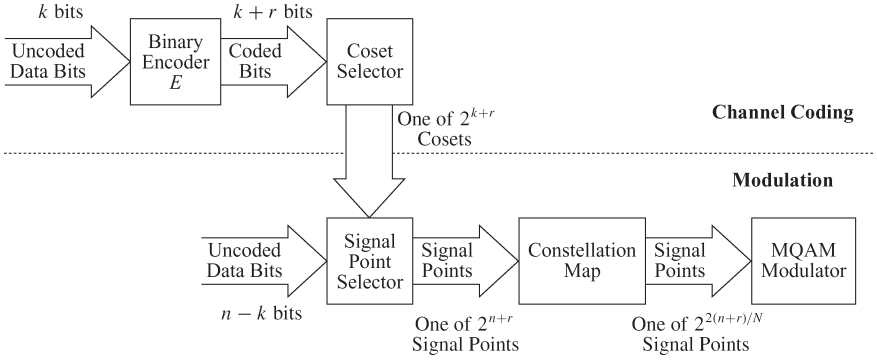


Figure 8.22: General coding scheme.

There are five elements required to generate the coded modulation:

1. a binary encoder E, block or convolutional, that operates on k uncoded data bits to produce $k+r$ coded bits;
2. a subset (coset) selector, which uses the coded bits to choose one of 2^{k+r} subsets from a partition of the N-dimensional signal constellation;
3. a point selector, which uses $n-k$ additional uncoded bits to choose one of the 2^{n-k} signal points in the selected subset;
4. a constellation map, which maps the selected point from N-dimensional space to a sequence of $N/2$ points in two-dimensional space; and
5. an MQAM modulator (or other M-ary modulator).

The first two elements constitute the channel coding, and the remaining elements are the modulation. The receiver essentially reverses the modulation and coding steps. After MQAM demodulation and an inverse $2/N$ constellation mapping, decoding is done in essentially two stages: first, the points within each subset that are closest to the received signal point are determined; then, the maximum likelihood subset sequence is calculated. When the encoder E is a convolutional encoder, this scheme is referred to as *trellis coded modulation*; for E a block encoder, it is called *lattice coded modulation*. These schemes are also referred to as trellis or lattice codes.

Steps 1–5 essentially decouple the channel coding gain from the gain associated with signal shaping in the modulation. Specifically, the code distance properties – and thus the channel coding gain – are determined by the encoder (E) properties and the subset partitioning, which are essentially decoupled from signal shaping. We will discuss the channel coding gain in more detail below. Optimal shaping of the signal constellation provides up to an additional 1.53 dB of shaping gain (for asymptotically large N), independent of the channel coding scheme.[1] However, the performance improvement from shaping gain is offset by the corresponding complexity of the constellation map, which grows exponentially with N. The size of the transmit constellation is determined by the average power constraint and does not affect the shaping or coding gain.

The channel coding gain results from a selection of sequences among all possible sequences of signal points. If we consider a sequence of N input bits as a point in N-dimensional space (the *sequence space*), then this selection is used to guarantee some minimum distance $d_{\min}$ in the sequence space between possible input sequences. Errors generally occur when a sequence is mistaken for its closest neighbor, and in AWGN channels this error probability is a decreasing function of $d^2_{\min}$. We can thus decrease the BER by increasing the separation between each point in the sequence space by a fixed amount ("stretching" the space). However, this will result in a proportional power increase, so no net coding gain is realized. The effective power gain of the channel code is therefore the minimum squared distance between selected sequence points (the sequence points obtained through coding) multiplied by the density of the selected sequence points. Specifically, if the minimum distance and density of all points in the sequence space are denoted by d_0 and $\triangle_0$, respectively, and if the minimum distance and density of points in the sequence space selected through coding are denoted by $d_{\min}$ and $\triangle$, respectively, then maximum likelihood sequence detection yields a channel coding gain of

$$G_c = \left(\frac{d^2_{\min}}{d^2_0}\right)\left(\frac{\triangle}{\triangle_0}\right). \tag{8.94}$$

The second term on the right side of this expression is also referred to as the *constellation expansion factor* and equals 2^{-r} (per N dimensions) for a redundancy of r bits in the encoder E [44].

Some of the nominal coding gain in (8.94) is lost owing to selected sequences having more than one nearest neighbor in the sequence space, which increases the possibility of incorrect sequence detection. This loss in coding gain is characterized by the *error coefficient*, which is tabulated for most common lattice and trellis coded modulations in [44]. In general, the error coefficient is larger for lattice codes than for trellis codes with comparable values of G_c.

Channel coding is done using set partitioning of lattices. A *lattice* is a discrete set of vectors in real Euclidean N-dimensional space that forms a group under ordinary vector addition, so the sum or difference of any two vectors in the lattice is also in the lattice. A *sublattice* is a subset of a lattice that is itself a lattice. The sequence space for *uncoded* MQAM modulation is just the N-cube,[2] so the minimum distance between points is no different than in the two-dimensional case. By restricting input sequences to lie on a lattice in N-dimensional space that is denser than the N-cube, we can increase $d_{\min}$ while maintaining the same density (or, equivalently, the same average power) in the transmit signal constellation; hence, there is no constellation expansion. The N-cube is a lattice, but

[1] A square constellation has 0 dB of shaping gain; a circular constellation, which is the geometrical figure with the least average energy for a given area, achieves the maximum shape gain for a given N [44].

[2] The Cartesian product of two-dimensional rectangular lattices with points at odd integers.

for every $N > 1$ there are denser lattices in N-dimensional space. Finding the densest lattice in N dimensions is a well-known mathematical problem, and it has been solved for all N for which the decoder complexity is manageable.[3] Once the densest lattice is known, we can form partitioning subsets, or *cosets*, of the lattice via translation of any sublattice. The choice of the partitioning sublattice will determine the size of the partition – that is, the number of subsets that the subset selector in Figure 8.22 has to choose from. Data bits are then conveyed in two ways: through the sequence of cosets from which constellation points are selected, and through the points selected within each coset. The density of the lattice determines the distance between points within a coset, while the distance between subset sequences is essentially determined by the binary code properties of the encoder E and its redundancy r. If we let d_p denote the minimum distance between points within a coset and d_s the minimum distance between the coset sequences, then the minimum distance is $d_{\min} = \min(d_p, d_s)$. The effective coding gain is given by

$$G_c = 2^{-2r/N} d_{\min}^2, \tag{8.95}$$

where $2^{-2r/N}$ is the constellation expansion factor (in two dimensions) from the r extra bits introduced by the binary channel encoder.

Returning to Figure 8.22, suppose we want to send $m = n + r$ bits per dimension, so that an N sequence conveys mN bits. If we use the densest lattice in N-dimensional space that lies within an N-dimensional sphere, where the radius of the sphere is just large enough to enclose 2^{mN} points, then we achieve a total coding gain that combines the coding gain (resulting from the lattice density and the encoder properties) with the shaping gain of the N-dimensional sphere over the N-dimensional rectangle. Clearly, the coding gain is decoupled from the shaping gain. An increase in signal power would allow us to use a larger N-dimensional sphere and hence transmit more uncoded bits. It is possible to generate maximum-density N-dimensional lattices for $N = 4, 8, 16$, and 24 using a simple partition of the two-dimensional rectangular lattice combined with either conventional block or convolutional coding. Details of this type of code construction, and the corresponding decoding algorithms, can be

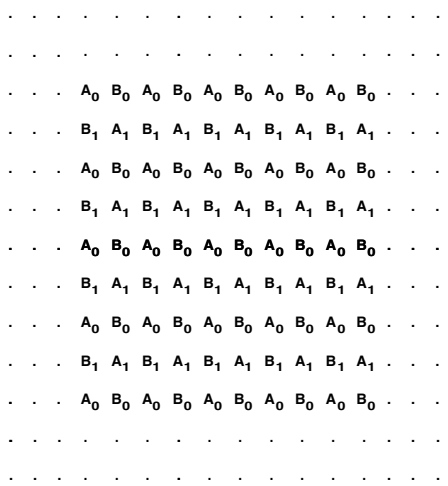


Figure 8.23: Subset partition for an 8-dimensional lattice.

found in [46] for both lattice and trellis codes. For these constructions, an effective coding gain of approximately 1.5, 3.0, 4.5, and 6.0 dB is obtained with lattice codes for $N = 4, 8, 16$, and 24, respectively. Trellis codes exhibit higher coding gains with comparable complexity.

[3]The complexity of the maximum likelihood decoder implemented with the Viterbi algorithm is roughly proportional to N.

We conclude this section with an example of coded modulation: the $N = 8$, 3dB–gain lattice code proposed in [46]. First, the two-dimensional signal constellation is partitioned into four subsets as shown in Figure 8.23, where the subsets are represented by the points A_0, A_1, B_0, and B_1. For example, a 16-QAM constellation would have four subsets, each consisting of four constellation points. Note that the distance between points in each subset is twice the distance between points in the (uncoded) constellation. From this subset partition, we form an 8-dimensional lattice by taking all sequences of four points in which (i) all points are either A-points or B-points and (ii) within a four-point sequence, the point subscripts satisfy the parity check $i_1 + i_2 + i_3 + i_4 = 0$ (so the sequence subscripts must be codewords in the (4, 3) parity-check code, which has a minimum Hamming distance of 2). Thus, three data bits and one parity check bit are used to determine the lattice subset. The square of the minimum distance resulting from this subset partition is four times that of the uncoded signal constellation, yielding a 6dB gain. However, the extra parity check bit expands the constellation by 1/2 bit per dimension, which by Section 5.3.3 costs an additional power factor of $4^{.5} = 2$, or 3 dB. Thus, the net coding gain is $6 - 3 = 3\,\mathrm{dB}$. The remaining data bits are used to choose a point within the selected subset and so, for a data rate of m bits per symbol, the four lattice subsets must each have 2^{m-1} points.[4] For example, with a 16-QAM constellation the four subsets each have $2^{m-1} = 4$ points, so the data rate is 3 data bits per 16-QAM symbol.

Coded modulation using turbo codes has also been investigated [47, 48, 49]. This work shows that turbo trellis coded modulation can come very close to the Shannon limit for non-binary signaling.

8.10 Coding with Interleaving for Fading Channels

Block, convolutional, and coded modulation are designed for good performance in AWGN channels. In fading channels, errors associated with the demodulator tend to occur in bursts, corresponding to the times when the channel is in a deep fade. Most codes designed for AWGN channels cannot correct for the long bursts of errors exhibited in fading channels. Hence, codes designed for AWGN channels can exhibit worse performance in fading than an uncoded system.

To improve performance of coding in fading channels, coding is typically combined with *interleaving* to mitigate the effect of error bursts. The basic premise of coding and interleaving is to spread error bursts due to deep fades over many codewords so that each received codeword exhibits at most a few simultaneous symbol errors, which can be corrected for. The spreading out of burst errors is accomplished by the interleaver and the error correction is accomplished by the code. The size of the interleaver must be large enough that fading is independent across a received codeword. Slowly fading channels require large interleavers, which in turn can lead to large delays.

Coding and interleaving is a form of diversity, and performance of coding and interleaving is often characterized by the diversity order associated with the resulting probability of error. This diversity order is typically a function of the minimum Hamming distance of the code. Thus, designs for coding and interleaving on fading channels must focus on maximizing the diversity order of the code rather than on metrics like Euclidean distance, which are used as a performance criterion in AWGN channels. In the following sections we discuss coding and interleaving for block, convolutional, and coded modulation in more detail. We will assume that the receiver has knowledge of the channel fading, which greatly simplifies both the analysis and the decoder. Estimates of channel fading are commonly obtained through pilot symbol transmissions [50, 51]. Maximum likelihood detection of coded signals in fading without this channel knowledge is computationally intractable [52] and so usually requires approximations to either the ML decoding metric or the channel [52, 53, 54, 55, 56, 57]. Note that turbo codes designed for AWGN channels, described in Section 8.6, have an interleaver inherent to the code design. However, the interleaver design considerations for AWGN channels are different than for fading channels. A discussion of interleaver design and performance analysis for turbo codes in fading channels can be found in [21, Chap. 8],

[4] This yields $m - 1$ bits/symbol, with the additional bit/symbol conveyed by the channel code.

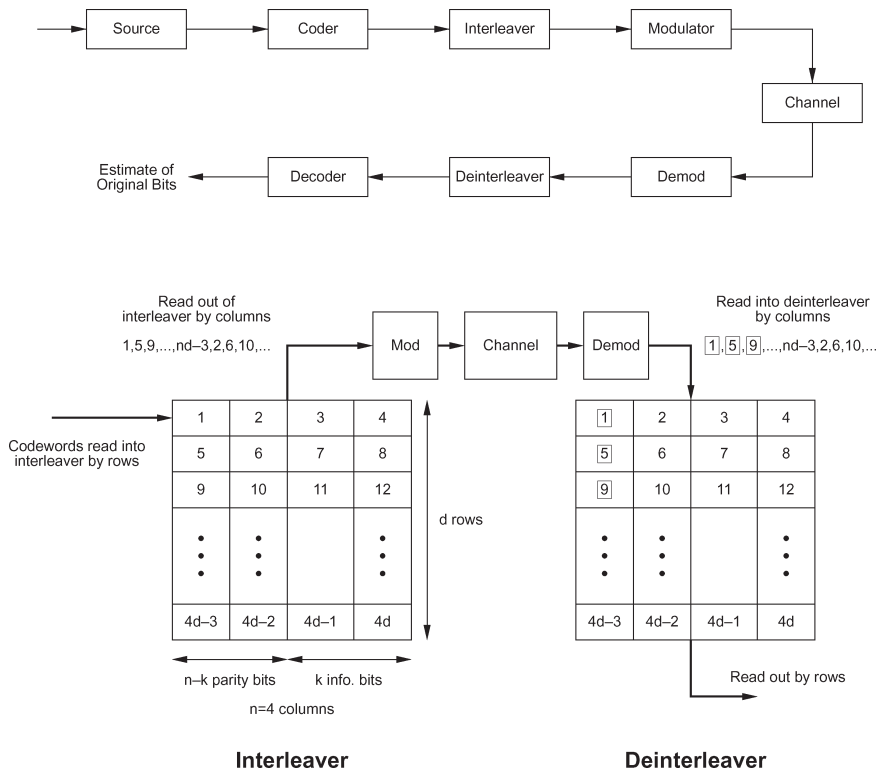


Figure 8.24: The interleaver/deinterleaver operation.

[58] and [59]. Low-density parity-check codes can also be designed for inherent diversity against fading, with performance similar to that of turbo codes optimized for fading diversity [60].

8.10.1 Block Coding with Interleaving

Block codes are typically combined with block interleaving to spread out burst errors from fading. A block interleaver is an array with d rows and n columns, as shown in Figure 8.24. For block interleavers designed for an (n, k) block code, codewords are read into the interleaver by rows so that each row contains an (n, k) codeword. The interleaver contents are read out by columns into the modulator for subsequent transmission over the channel. During transmission, codeword symbols in the same codeword are separated by $d - 1$ other symbols, so symbols in the same codeword experience approximately independent fading if their separation in time is greater than the channel coherence time – that is, if $dT_s > T_c \approx 1/B_D$, where T_s is the codeword symbol duration, T_c is the channel coherence time, and B_D is the channel Doppler spread. An interleaver is called a *deep interleaver* if the condition $dT_s > T_c$ is satisfied. The deinterleaver is an array identical to the interleaver. Bits are read into the deinterleaver from the demodulator by column so that each row of the deinterleaver contains a codeword (whose bits may have been corrupted by the channel). The deinterleaver output is read into the decoder by rows, one codeword at a time.

Figure 8.24 illustrates the ability of coding and interleaving to correct for bursts of errors. Suppose our coding scheme is an (n, k) binary block code with error correction capability $t = 2$. If this codeword is transmitted through a channel with an error burst of three symbols, then three out of four of the codeword symbols will be received in error. Since the code can only correct two or fewer errors, the codeword will be decoded in error. However, if the codeword is put through an interleaver then, as shown in Figure 8.24, the error burst of three symbols will be

spread out over three separate codewords. Since a single symbol error can be easily corrected by an (n, k) code with $t = 2$, the original information bits can be decoded without error. Convolutional interleavers are similar in concept to block interleavers and are better suited to convolutional codes, as will be discussed in Section 8.10.2.

Performance analysis of coding and interleaving requires pairwise error probability analysis or the application of Chernoff or union bounds. Details of this analysis can be found in [6, Chap. 14.6]. The union bound provides a simple approximation to performance. Assume a Rayleigh fading channel with deep interleaving such that each coded symbol fades independently. Then the union bound for an (n, k) block code with soft decision decoding under noncoherent FSK modulation yields a codeword error given as

$$P_e < (2^k - 1)[4p(1-p)]^{d_{\min}}, \tag{8.96}$$

where $d_{\min}$ is the minimum Hamming distance of the code and

$$p = \frac{1}{2 + R_c \bar{\gamma}_b}. \tag{8.97}$$

Similarly, for slowly fading channels in which a coherent phase reference can be obtained, the union bound on the codeword error probability of an (n, k) block code with soft decision decoding and BPSK modulation yields

$$P_e < 2^k \binom{2d_{\min} - 1}{d_{\min}} \left(\frac{1}{4R_c \bar{\gamma}_b} \right)^{d_{\min}}. \tag{8.98}$$

Note that both (8.96) and (8.98) are similar to the formula for error probability under MRC diversity combining given by (7.34), with $d_{\min}$ providing the diversity order. Similar formulas apply for hard decoding, with diversity order reduced by a factor of 2 relative to soft decision decoding. Thus, designs for block coding and interleaving over fading channels optimize their performance by maximizing the Hamming distance of the code.

Coding and interleaving is a suboptimal coding technique, since the correlation of the fading that affects subsequent bits contains information about the channel that could be used in a true maximum likelihood decoding scheme. By essentially throwing away this information, the inherent capacity of the channel is decreased [61]. Despite this capacity loss, coding with interleaving using codes designed for AWGN channels is a common coding technique for fading channels, since the complexity required for maximum likelihood decoding on correlated coded symbols is prohibitive.

Example 8.8: Consider a Rayleigh fading channel with a Doppler of $B_D = 80\,\text{Hz}$. The system uses a (5,2) Hamming code with interleaving to compensate for the fading. If the codeword symbols are sent through the channel at 30 kbps, find the required interleaver depth needed to obtain independent fading on each symbol. What is the longest burst of codeword symbol errors that can be corrected and the total interleaver delay for this depth?
Solution: The (5, 2) Hamming code has a minimum distance of 3, so it can correct $t = \lfloor .5 \cdot 3 \rfloor = 1$ codeword symbol error. The codeword symbols are sent through the channel at a rate $R_s = 30$ kbps, so the symbol time is $T_s = 1/R_s = 3.3 \cdot 10^{-5}$. Assume a coherence time for the channel of $T_c = 1/B_D = .0125$ s. The bits in the interleaver are separated by dT_s, so we require $dT_s \geq T_c$ for independent fading on each codeword symbol. Solving for d yields $d \geq T_c/T_s = 375$. Since the interleaver spreads a burst of errors over the depth d of the interleaver, a burst of d symbol errors in the interleaved codewords will result in just one symbol error per codeword after deinterleaving, which can be corrected. The system can therefore tolerate an error burst of 375 symbols. However, all rows of the interleaver must be filled before it can read out by columns, so the total delay of the interleaver is $ndT_s = 5 \cdot 375 \cdot 3.3 \cdot 10^{-5} = 62.5$ ms. This delay can degrade quality in a voice system. We thus see that the price paid for correcting long error bursts through coding and interleaving is significant delay.

8.10.2 Convolutional Coding with Interleaving

As with block codes, convolutional codes suffer performance degradation in fading channels because the code is not designed to correct for bursts of errors. Thus, it is common to use an interleaver to spread out error bursts. In block coding the interleaver spreads errors across different codewords. Since there is no similar notion of a codeword in convolutional codes, a slightly different interleaver design is needed to mitigate the effect of burst errors. The interleaver commonly used with convolutional codes, called a *convolutional interleaver*, is designed both to spread out burst errors and to work well with the incremental nature of convolutional code generation [62, 63].

A block diagram for a convolutional interleaver is shown in Figure 8.25. The encoder output is multiplexed into buffers of increasing size, from no buffering to a buffer of size $N-1$. The channel input is similarly multiplexed from these buffers into the channel. The reverse operation is performed at the decoder. Thus, the convolutional interleaver delays the transmission through the channel of the encoder output by progressively larger amounts, and this delay schedule is reversed at the receiver. This interleaver takes sequential outputs of the encoder and separates them by $N-1$ other symbols in the channel transmission, thereby breaking up burst errors in the channel. Note that a convolutional encoder can also be used with a block code, but it is most commonly used with a convolutional code. The total memory associated with the convolutional interleaver is $.5N(N-1)$ and the delay is $N(N-1)T_s$ [7, Chap. 8.2.2], where T_s is the symbol time for transmitting the coded symbols over the channel.

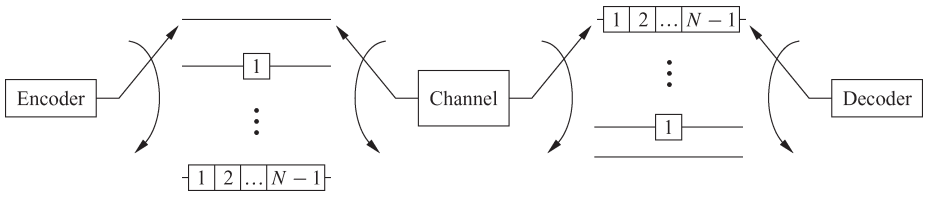


Figure 8.25: Convolutional coding and interleaving.

The probability of error analysis for convolutional coding and interleaving is given in [6, Chap. 14.6] under assumptions similar to those used in our block fading analysis. The Chernoff bound again yields a probability of error under soft decision decoding with a diversity order based on the minimum free distance of the code. Hard decision decoding reduces this diversity by a factor of 2.

Example 8.9: Consider a channel with coherence time $T_c = 12.5$ ms and a coded bit rate of $R_s = 100$ kilosymbols per second. Find the average delay of a convolutional interleaver that achieves independent fading between subsequent coded bits.

Solution: For the convolutional interleaver, each subsequent coded bit is separated by NT_s and we require $NT_s \geq T_c$ for independent fading, where $T_s = 1/R_s$. Thus we have $N \geq T_c/T_s = .0125/.00001 = 1250$. Note that this is the same as the required depth for a block interleaver to get independent fading on each coded bit. The total delay is $N(N-1)T_s = 15$ s. This is quite a high delay for either voice or data.

8.10.3 Coded Modulation with Symbol/Bit Interleaving

As with block and convolutional codes, coded modulation designed for an AWGN channel performs poorly in fading. This leads to the notion of coded modulation with interleaving for fading channels. However, unlike block and convolutional codes, there are two options for interleaving in coded modulation. One option is to interleave the bits and then map them to modulated symbols. This is called bit-interleaved coded modulation (BICM).

Alternatively, the modulation and coding can be done jointly as in coded modulation for AWGN channels and the resulting symbols interleaved prior to transmission. This technique is called symbol-interleaved coded modulation (SICM).

Symbol-interleaved coded modulation seems at first like the best approach because it preserves joint coding and modulation, the main design premise behind coded modulation. However, the coded modulation design criterion must be changed in fading, since performance in fading depends on the code diversity as characterized by its Hamming distance rather than its Euclidean distance. Initial work on coded modulation for fading channels focused on techniques to maximize diversity in SICM. However, good design criteria were hard to obtain, and the performance of these codes was somewhat disappointing [64, 65, 66].

A major breakthrough in the design of coded modulation for fading channels was the discovery of bit-interleaved coded modulation [67, 68]. In BICM the code diversity is equal to the smallest number of distinct bits (rather than channel symbols) along any error event. This is achieved by bitwise interleaving at the encoder output prior to symbol mapping, with an appropriate soft decision bit metric as an input to the Viterbi decoder. Although this breaks the coded modulation paradigm of joint modulation and coding, it provides much better performance than SICM. Moreover, analytical tools for evaluating the performance of BICM as well as design guidelines for good performance are known [67]. BICM is now the dominant technique for coded modulation in fading channels.

8.11 Unequal Error Protection Codes

When not all bits transmitted over the channel have the same priority or bit error probability requirement, multiresolution or unequal error protection (UEP) codes can be used. This scenario arises, for example, in voice and data systems where voice is typically more tolerant to bit errors than data: data received in error must be retransmitted, so $P_b < 10^{-6}$ is typically required, whereas good quality voice requires only on the order of $P_b < 10^{-3}$. This scenario also arises for certain types of compression. For example, in image compression, bits corresponding to the low-resolution reproduction of the image are required, whereas high-resolution bits simply refine the image. With multiresolution channel coding, all bits are received correctly with a high probability under benign channel conditions. However, if the channel is in a deep fade, only the high-priority bits or those requiring low P_b will be received correctly with high probability.

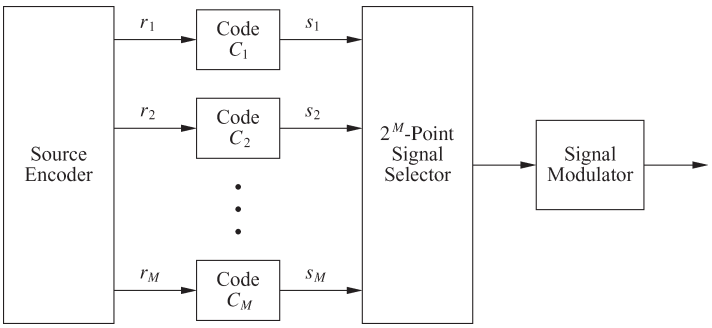


Figure 8.26: Multilevel encoder.

Practical implementation of a UEP code was first studied by Imai and Hirakawa [69]. Binary UEP codes were later considered for combined speech and channel coding [70] as well as for combined image and channel coding [71]. These implementations use traditional (block or convolutional) error correction codes, so coding gain is directly proportional to bandwidth expansion. Subsequently, two bandwidth-efficient implementations for UEP codes were proposed: time multiplexing of bandwidth-efficient coded modulation [72], and coded modulation

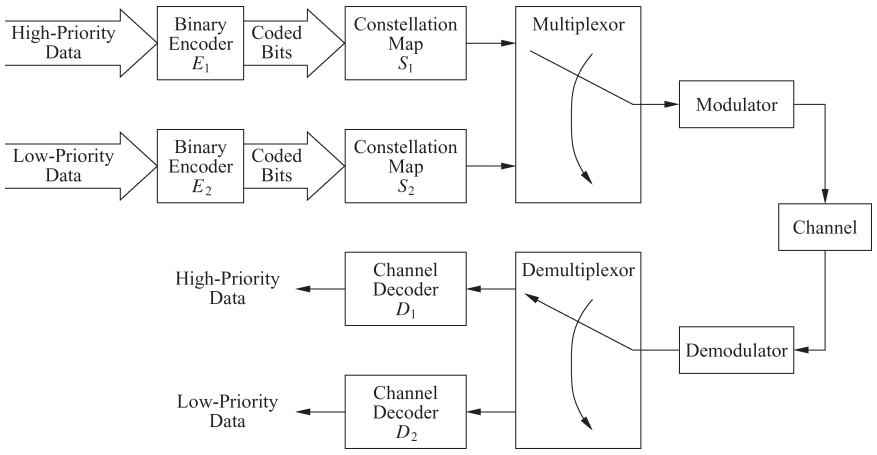


Figure 8.27: Transceiver for time-multiplexed coded modulation.

techniques applied to both uniform and nonuniform signal constellations [73, 74]. All of these multilevel codes can be designed for either AWGN or fading channels. We now briefly summarize these UEP coding techniques; specifically, we describe the principles behind multilevel coding and multistate decoding as well as the more complex bandwidth-efficient implementations.

A block diagram of a general multilevel encoder is shown in Figure 8.26. The source encoder first divides the information sequence into M parallel bit streams of decreasing priority. The channel encoder consists of M different binary error correcting codes $C_1, \ldots, C_M$ with decreasing codeword distances. The ith-priority bit stream enters the ith encoder, which generates the coded bits s_i. If the 2^M points in the signal constellation are numbered from 0 to $2^M - 1$, then the point selector chooses the constellation point s corresponding to

$$s = \sum_{i=1}^{M} s_i \cdot 2^{i-1}. \tag{8.99}$$

For example, if $M = 3$ and the signal constellation is 8-PSK, then the chosen signal point will have phase $2\pi s/8$.

Optimal decoding of the multilevel code uses a maximum likelihood decoder, which determines the input sequence that maximizes the received sequence probability. The ML decoder must therefore jointly decode the code sequences $\{s_1\}, \ldots, \{s_m\}$. This can entail significant complexity even if the individual codes in the multilevel code have low complexity. For example, if the component codes are convolutional codes with 2^{μ_i} states, $i = 1, \ldots, M$, then the number of states in the optimal decoder is $2^{\mu_1 + \cdots + \mu_M}$. Because of the high complexity of optimal decoding, the suboptimal technique of multistage decoding, introduced in [69], is used for most implementations. Multistage decoding is accomplished by decoding the component codes sequentially. First the most powerful code, C_1, is decoded, then C_2, and so forth. Once the code sequence corresponding to encoder C_i is estimated, it is assumed to be correct for code decisions on the weaker code sequences.

The binary encoders of this multilevel code require extra code bits to achieve their coding gain, so they are not bandwidth efficient. An alternative approach proposed in [73] uses time multiplexing of the coded modulations described in Section 8.9. In this approach, different lattice or trellis coded modulations with different coding gains are used for each priority class of input data. The transmit signal constellations corresponding to each encoder may differ in size (number of signal points), but the average power of each constellation is the same. The signal points output by each of the individual encoders are then time-multiplexed together for transmission over the channel, as shown in Figure 8.27 for two bit streams of different priority. Let R_i denote the bit rate of encoder C_i in this figure for $i = 1, 2$. If T_1 equals the fraction of time that the high-priority C_1 code is transmitted and if T_2 equals

the fraction of time that the C_2 code is transmitted, then the total bit rate is $(R_1T_1 + R_2T_2)/(T_1 + T_2)$, with the high-priority bits constituting $R_1T_1/(R_1T_1 + R_2T_2)$ percent of this total.

The time-multiplexed coding method yields a higher gain if the constellation maps S_1 and S_2 of Figure 8.27 are designed jointly. This revised scheme is shown in Figure 8.28 for two encoders, where the extension to M encoders is straightforward. In trellis and lattice coded modulation, recall that bits are encoded to select the lattice subset and that uncoded bits choose the constellation point within the subset. The binary encoder properties reduce the P_b for the encoded bits only; the P_b for the uncoded bits is determined by the separation of the constellation signal points. We can easily modify this scheme to yield two levels of coding gain, where the high-priority bits are heavily encoded and are used to choose the subset of the partitioned constellation, while the low-priority bits are uncoded or lightly coded and are used to select the constellation signal point.

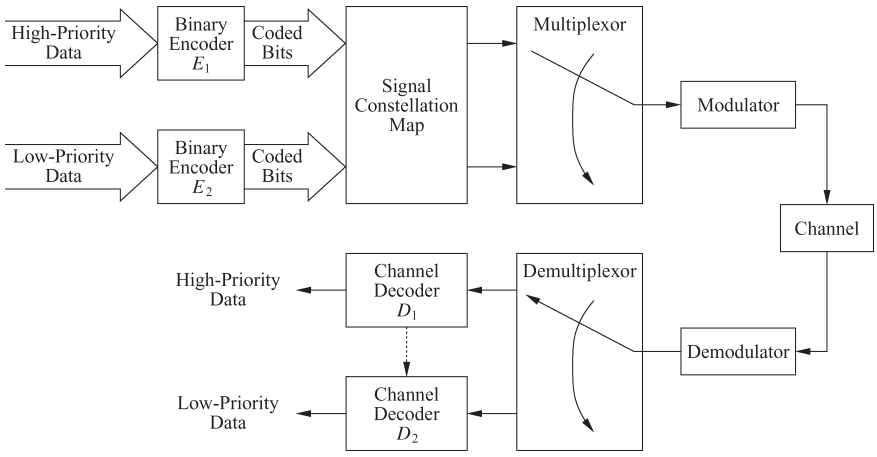


Figure 8.28: Joint optimization of signal constellation.

8.12 Joint Source and Channel Coding

The underlying premise of UEP codes is that the bit error probabilities of the channel code should be matched to the priority or P_b requirements associated with the bits to be transmitted. These bits are often taken from the output of a compression algorithm acting on the original data source. Hence, UEP coding can be considered as a joint design between compression (also called *source coding*) and channel coding. Although Shannon determined that the source and channel codes can be designed separately on an AWGN channel with no loss in optimality [75], this result holds only in the limit of infinite source code dimension, infinite channel code blocklength, and infinite complexity and delay. Thus, there has been much work on investigating the benefits of joint source and channel coding under more realistic system assumptions.

Work in the area of joint source and channel coding falls into several broad categories: source-optimized channel coding; channel-optimized source coding; and iterative algorithms, which combine these two code designs. In source-optimized channel coding, the source code is designed for a noiseless channel. A channel code is then designed for this source code to minimize end-to-end distortion over the given channel based on the distortion associated with corruption of the different transmitted bits. Unequal error protection channel coding – where the P_b of the different component channel codes is matched to the bit priorities associated with the source code – is an example of this technique. Source-optimized channel coding has been applied to image compression with convolution channel coding and with rate-compatible punctured convolutional (RCPC) channel codes in [71, 76, 77]. A comprehensive treatment of matching RCPC channel codes or MQAM to subband and linear predictive speech

coding – in both AWGN and Rayleigh fading channels – can be found in [78]. In source-optimized modulation, the source code is designed for a noiseless channel and then the modulation is optimized to minimize end-to-end distortion. An example of this approach is given in [79], where compression by a vector quantizer (VQ) is followed by multicarrier modulation, and the modulation provides unequal error protection to the different source bits by assigning different powers to each subcarrier.

Channel-optimized source coding is another approach to joint source and channel coding. In this technique the source code is optimized based on the error probability associated with the channel code, where the channel code is designed independently of the source. Examples of work taking this approach include the channel-optimized vector quantizer (COVQ) and its scalar variation [80, 81]. Source-optimized channel coding and modulation can be combined with channel-optimized source coding via an iterative design. This approach is used for the joint design of a COVQ and multicarrier modulation in [82] and for the joint design of a COVQ and RCPC channel code in [83]. Combined trellis coded modulation and *trellis coded quantization*, a source coding strategy that borrows from the basic premise of trellis coded modulation, is investigated in [84, 85]. All of this work on joint source and channel code design indicates that significant performance advantages are possible when the source and channel codes are jointly designed. Moreover, many sophisticated channel code designs, such as turbo and LDPC codes, have not yet been combined with source codes in a joint optimization. Thus, much more work is needed in the broad area of joint source and channel coding to optimize performance for different applications.

Chapter 8 Problems

1. Consider a $(3, 1)$ linear block code where each codeword consists of three data bits and one parity bit.

   (a) Find all codewords in this code.

   (b) Find the minimum distance of the code.

2. Consider a $(7, 4)$ code with generator matrix

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

   (a) Find all the codewords of the code.

   (b) What is the minimum distance of the code?

   (c) Find the parity check matrix of the code.

   (d) Find the syndrome for the received vector $\mathbf{R} = [1101011]$.

   (e) Assuming an information bit sequence of all 0s, find all minimum weight error patterns $\mathbf{e}$ that result in a valid codeword that is not the all-zero codeword.

   (f) Use row and column operations to reduce $\mathbf{G}$ to systematic form and find its corresponding parity-check matrix. Sketch a shift register implementation of this systematic code.

3. All Hamming codes have a minimum distance of 3. What are the error correction and error detection capabilities of a Hamming code?

4. The $(15, 11)$ Hamming code has generator polynomial $g(X) = 1 + X + X^4$. Determine if the codewords described by polynomials $c_1(X) = 1 + X + X^3 + X^7$ and $c_2(X) = 1 + X^3 + X^5 + X^6$ are valid codewords for this generator polynomial. Also find the systematic form of this polynomial $p(X) + X^{n-k}u(X)$ that generates the codewords in systematic form.

5. The $(7, 4)$ cyclic Hamming code has a generator polynomial $g(X) = 1 + X^2 + X^3$.

   (a) Find the generator matrix for this code in systematic form.

   (b) Find the parity-check matrix for the code.

   (c) Suppose the codeword $\mathbf{C} = [1011010]$ is transmitted through a channel and the corresponding received codeword is $\mathbf{C} = [1010011]$. Find the syndrome polynomial associated with this received codeword.

   (d) Find all possible received codewords such that, for the transmitted codeword $\mathbf{C} = [1011010]$, the received codeword has a syndrome polynomial of zero.

6. Derive (8.52), i.e. show that the probability of mistaking the all-zero codeword $\mathbf{C_0}$ for another codeword $\mathbf{C_i}$ with Hamming weight w_i under BPSK transmission over an AWGN channel with energy per coded symbol E_c and SDD is given by

$$P_e(\mathbf{C}_i) = Q\left(\frac{2w_i\sqrt{E_c}}{\sqrt{2w_i N_0}}\right).$$

7. The weight distribution of a Hamming code of blocklength n is given by

$$N(x) = \sum_{i=0}^{n} N_i x^i = \frac{1}{n+1}[(1+x)^n + n(1+x)^{.5(n-1)}(1-x)^{.5(n+1)}],$$

where N_i denotes the number of codewords of weight i.

   (a) Use this formula to determine the weight distribution of a Hamming $(7, 4)$ code.

   (b) Use the weight distribution from part (a) to find the union upper bound (8.53) based on weight distribution (8.48) for a Hamming $(7, 4)$ code, assuming BPSK modulation of the coded bits with an SNR of 10 dB. Compare with the probability of error from the looser bound (8.49) for the same modulation.

8. Find the union upper bound (8.53) on probability of codeword error for a Hamming code with $m = 7$. Assume the coded bits are transmitted over an AWGN channel using 8-PSK modulation with an SNR of 10 dB. Compute the probability of bit error for the code assuming a codeword error corresponds to one bit error, and compare with the bit error probability for uncoded modulation.

9. Plot P_b versus γ_b for a $(24, 12)$ linear block code with $d_{\min} = 8$ and $0 \leq E_b/N_0 \leq 20$ dB using the union bound for probability of codeword error. Assume that the coded bits are transmitted over the channel using QPSK modulation. Over what range of E_b/N_0 does the code exhibit negative coding gain?

10. Use (8.58) to find the approximate coding gain of a $(7, 4)$ Hamming code with SDD over uncoded modulation, assuming $\gamma_b = 15$ dB.

11. Plot the probability of codeword error for a $(24, 12)$ code with $d_{\min} = 8$ for $0 \leq \gamma_b \leq 10$ dB under both hard and soft decoding, using the union bound (8.46) for hard decoding and the approximation (8.55) for soft decoding. What is the difference in coding gain at high SNR for the two decoding techniques?

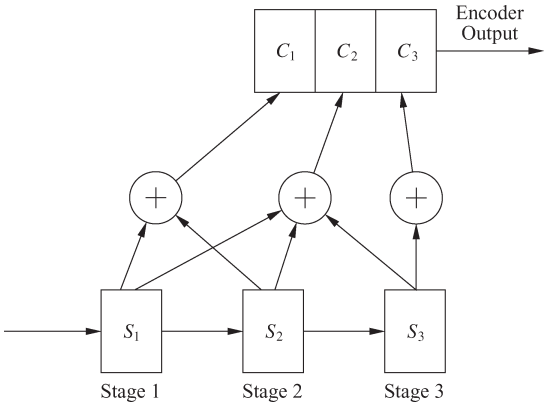


Figure 8.29: Convolutional encoder for Problems 8-14 and 8-15.

12. Evaluate the upper and lower bounds on codeword error probability, (8.45) and (8.46) respectively, for an extended Golay code with HDD, assuming an AWGN channel with BPSK modulation and an SNR of 10 dB.

13. Consider a Reed Solomon code with $k = 3$ and $K = 4$, mapping to 8-PSK modulation. Find the number of errors that can be corrected with this code and its minimum distance. Also find its probability of bit error assuming the coded symbols transmitted over the channel via 8-PSK have a symbol error probability of 10^{-3}.

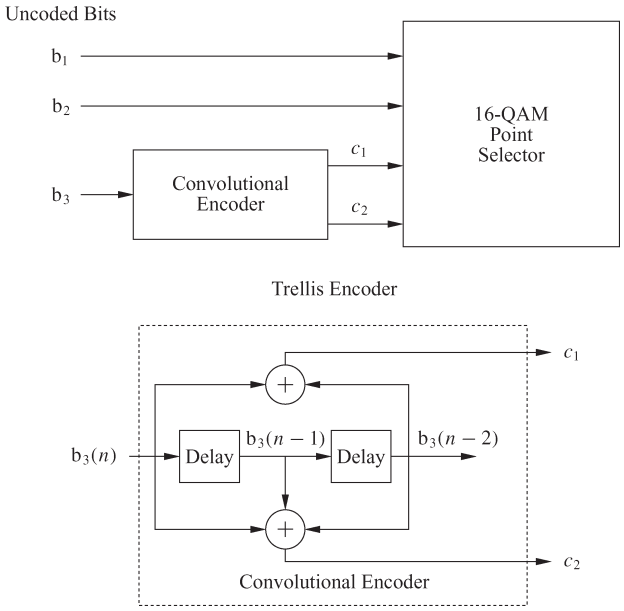


Figure 8.30: 16-QAM trellis encoder for Problem 8-16.

14. For the trellis of Figure 8.9, determine the state sequence and encoder output assuming an initial state $S = 00$ and information bit sequence $\mathbf{U} = [0110101101]$.

15. Consider the convolutional code generated by the encoder shown in Figure 8.29.

    (a) Sketch the trellis diagram of the code.

    (b) For a received sequence $\mathbf{R} = [001010001]$, find the path metric for the all-zero path assuming probability of symbol error $p = 10^{-3}$.

    (c) Find one path at a minimum Hamming distance from the all-zero path and compute its path metric for the same $\mathbf{R}$ and p as in part (b).

16. This problem is based on the convolutional encoder of Figure 8.29.

    (a) Draw the state diagram for this convolutional encoder.

    (b) Determine its transfer function $T(D,\ N,\ J)$.

    (c) Determine the minimum distance of paths through the trellis to the all-zero path.

    (d) Compute the upper bound (8.86) on probability of bit error for this code assuming SDD and BPSK modulation with $\gamma_b = 10$ dB.

    (e) Compute the upper bound (8.90) on probability of bit error for this code assuming HDD and BPSK modulation with $\gamma_b = 10$ dB. How much coding gain is achieved with soft versus hard decoding?

17. For the LDPC code with parity check matrix shown below, determine the code's Tanner Graph and its girth.

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{8.100}$$

18. For a polar code designed for a binary erasure channel W, find the mutual information for $W_8^i, i = 1, \ldots, 8$ and evaluate it for $\epsilon = .4$. How many of these channels have mutual information greater than $I(W)$?

19. Suppose you have a 16-QAM signal constellation that is trellis encoded via the scheme of Figure 8.30. Assume the set partitioning for 16-QAM shown in Figure 8.23.

    (a) Assuming that parallel transitions–that is, transitions within a subset–dominate the error probability, find the coding gain of this trellis code relative to uncoded 8-PSK, given that d_0 for the 16-QAM is .632 and for the 8-PSK is .765.

    (b) Draw the trellis for the convolutional encoder and assign subsets to the trellis transitions according to the following heuristic rules of Ungerboeck: (i) if k bits are to be encoded per modulation interval then the trellis must allow for $2k$ possible transitions from each state to the next state; (ii) more than one transition may occur between pairs of states; (iii) all waveforms should occur with equal frequency and with a fair amount of regularity and symmetry; (iv) transitions originating from the same state are assigned waveforms either from the A or B subsets but not from both; (v) transitions entering into the same state are assigned waveforms either from the A or B subsets but not from both; and (vi) parallel transitions are assigned from only one subset.

    (c) What is the minimum distance error event through the trellis relative to the path generated by the all-zero bit stream?

    (d) Assuming that your answer to part (c) is the minimum distance error event for the trellis, what is $d_{\min}$ of the code?

20. Consider a channel with coherence time $T_c = 10$ ms and a coded bit rate of $R_s = 50$ kilosymbols per second. Find the average delay of a convolutional interleaver that achieves independent fading between subsequent coded bits. Also find the memory requirements of this system.

21. In a Rayleigh fading channel, determine an upper bound for the bit error probability P_b of a Golay $(23, 12)$ code with deep interleaving $(dT_s \gg T_c)$, BPSK modulation, soft decision decoding, and an average coded E_c/N_0 of 15 dB. Compare with the uncoded P_b in Rayleigh fading.

22. Consider a Rayleigh fading channel with BPSK modulation, average SNR of 10 dB, and a Doppler of 80 Hz. The data rate over the channel is 30 kbps. Assume that bit errors occur on this channel whenever $P_b(\gamma) \geq 10^{-2}$. Design an interleaver and associated $(n,\ k)$ block code that corrects essentially all of the bit errors, where the interleaver delay is constrained to be less than 50 ms. Your design should include the dimensions of the interleaver as well as the block code type and the values of n and k.

23. Assume a multilevel encoder as in Figure 8.26, where the information bits have three different error protection levels ($M = 3$) and the three encoder outputs are modulated using 8-PSK modulation with an SNR of 10 dB. Assume the code C_i associated with the ith bit stream b_i is a Hamming code with parameter m_i, where $m_1 = 2, m_2 = 3$, and $m_3 = 4$.

    (a) Find the probability of error for each Hamming code C_i, assuming it is decoded individually using HDD.

    (b) If the symbol time of the 8-PSK modulation is $T_s = 10$ μs, what is the data rate for each of the three bit streams?

    (c) For what size code must the maximum likelihood decoder of this UEP code be designed?

24. Design a two-level UEP code using either Hamming or Golay codes and such that, for a channel with an SNR of 10 dB, the UEP code has $P_b = 10^{-3}$ for the low-priority bits and $P_b = 10^{-6}$ for the high priority bits.

Bibliography

[1] S. Lin and J. D. J. Costello, *Error Control Coding*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 2004.

[2] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[3] D. Declercq, M. Fossorier, E. Biglieri, *Channel Coding: Theory, Algorithms, and Applications,* Academic Press Library in Mobile and Wireless Communications, 2016.

[4] W. Ryan, and S. Lin, *Channel Codes: Classical and Modern* Cambridge, University Press, 2009.

[5] D. J. Costello and G. D. Forney, Jr., "Channel coding: The road to channel capacity," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150-1177, June 2007.

[6] J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001.

[7] B. Sklar, *Digital Communications–Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[8] S. G. Wilson, *Digital Modulation and Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1996.

[9] G. D. Forney, Jr., "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268-278, March 1973

[10] A. J. Viterbi, "Error bounds for convolutional codes and asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, pp. 260–9, April 1967.

[11] P. Elias, "Coding for noisy channels," *IRE Conv. Rec.,* Part 4, pp. 37–46, March 1955.

[12] G.D. Forney, Jr., "Concatenated Codes," *Technical Report 440, MIT Research Laboratory of Electronics*, December 1965. See also *Concatenated Codes*, MIT Press, Cambridge MA, 1966.

[13] P.A. Regalia, "Iterative Decoding of Concatenated Codes: A Tutorial." *EURASIP Journal on Advances in Signal Processing*, pp. 1–13, 2005.

[14] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," *Proc. IEEE Intern. Conf. Commun.*, pp. 54–8, May 1993.

[15] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, pp. 1261–71, October 1996.

[16] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, pp. 21–8, January 1962.

[17] S. B. Wicker and S. Kim, *Codes, Graphs, and Iterative Decoding*, Kluwer, Boston, 2002.

[18] *IEEE Trans. Inform. Theory*, Special Issue on Codes and Graphs and Iterative Algorithms, February 2001.

[19] C. Heegard and S. B. Wicker, *Turbo Coding*, Kluwer, Boston, 1999.

[20] L. Hanzo, T.H. Liew, B.L. Yeap, R.Y.S. Tee, and S. X. Ng, *Turbo Coding, Turbo Equalisation and Space-Time Coding: EXIT-Chart-Aided Near-Capacity Designs for Wireless Channels* , Wiley-IEEE Press, 2nd Ed., 2011.

[21] B. Vucetic and J. Yuan, *Turbo Codes, Principles and Applications*, Springer USA, 2000.

[22] L.R. Bahl, J. Cocke E. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inf. Theory*, Vol. 20, pp. 284–287, March 1974.

[23] M. AlMahamdy and J. Dill, "Early termination of turbo decoding by identification of undecodable blocks," *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, March 2017.

[24] M.C. Valenti, "Turbo codes and iterative processing," *Proc. IEEE New Zealand Wireless Commun. Symp.*, Nov. 1998.

[25] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design and iterative decoding," *IEEE Trans. Inform. Theory*, pp. 909–26, May 1998.

[26] H. Jin and R. J. McEliece, "Coding theorems for turbo code ensembles," *IEEE Trans. Inform. Theory*, pp. 1451–61, June 2002.

[27] I. Sasan and S. Shamai, "Improved upper bounds on the ML decoding error probability of parallel and serial concatenated turbo codes via their ensemble distance spectrum," *IEEE Trans. Inform. Theory*, pp. 24–47, January 2000.

[28] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Elec. Lett*., p. 1645, August 1996.

[29] N. Wiberg, N.-A. Loeliger, and R. Kotter, "Codes and iterative decoding on general graphs," *Euro. Trans. Telecommun*., pp. 513–25, June 1995.

[30] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, v pp. 533–547, Sep. 1981

[31] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message passing decoding," *IEEE Trans. Inform. Theory*, pp. 599–618, February 2001.

[32] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low density parity-check codes," *IEEE Trans. Inform. Theory*, pp. 619–37, February 2001.

[33] R. McEliece, D. J. C. MacKay, and J.-F. Cheng, " Turbo decoding as an instance of Pearl's "belief propagation" algorithm," *IEEE J. Sel. Areas Commun*., pp. 140–52, February 1998.

[34] F. R. Kschischang and D. Frey, "Iterative decoding of compound codes by probability propagation in graphical models," *IEEE J. Sel. Areas Commun*., pp. 219–30, February 1998.

[35] S.-Y. Chung, G. D. Forney, Jr., T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett*., pp. 58–60, Febmary 2001.

[36] M. Fossorier, "Iterative reliability-based decoding of low-density parity check codes," *IEEE J. Sel. Areas Commun*., pp. 908–17, May 2001.

[37] M. C. Davey and D. MacKay, "Low density parity-check codes over GF(q)," *IEEE Commun. Lett*., pp. 165–7, June 1998.

[38] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051-3073, July 2009.

[39] J. Hou, P. Siegel, L. Milstein, and H. D. Pfister, "Capacity-approaching bandwidth efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inform. Theory*, pp. 2141–55, September 2003.

[40] D. Hui, S. Sandberg, Y. Blankenship, M. Andersson, and L. Grosjean, "Channel Coding in 5G New Radio: A Tutorial Overview and Performance Comparison with 4G LTE," IEEE *Vehicular Technology Magazine*, vol. 13, no. 4, pp. 60-69, Dec. 2018.

[41] V. Bioglio, C. Condo and I. Land, "Design of polar codes in 5G new radio," *IEEE Communications Surveys and Tutorials*, vol. 23, no. 1, pp. 29-40, Jan. 2020.

[42] G. Ungerboeck, "Channel coding with multi-level/phase signals," *IEEE Trans. Inform. Theory*, pp. 55–67, January 1982.

[43] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets Part I: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5-11, February 1987.

[44] G. D. Forney, Jr., "Coset codes, I: Introduction and geometrical classification *IEEE Trans. Inform. Theory*, pp. 1123–51, September 1988.

[45] G. D. Forney, Jr., "Coset codes, II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, pp. 1152–87, September 1988.

[46] G. D. Forney, Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Quereshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun*., pp. 632–47, September 1984.

[47] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Parallel concatenated trellis coded modulation," *Proc. IEEE Internat. Conf. Commun*., pp. 974–8, June 1996.

[48] C. Fragouli and R. D. Wesel, "Turbo-encoder design for symbol-interleaved parallel concatenated trellis-coded modulation," *IEEE Trans. Commun*., pp. 425–35, March 2001.

[49] P. Robertson and T. Worz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. Sel. Areas Commun*., pp. 206–18, February 1998.

[50] G. T. Irvine and P. J. Mclane, "Symbol-aided plus decision-directed reception for PSK TCM modulation on shadowed mobile satellite fading channels," *IEEE J. Sel. Areas Commun*., pp. 1289–99, October 1992.

[51] D. Subasinghe-Dias and K. Feher, "A coded 16-QAM scheme for fast fading mobile radio channels," *IEEE Trans. Commun*., pp. 1906–16, February–April 1995.

[52] P. Y. Kam and H. M. Ching, "Sequence estimation over the slow nonselective Rayleigh fading channel with diversity reception and its application to Viterbi decoding," *IEEE J. Sel. Areas Commun*., pp. 562–70, April 1992.

[53] D. Makrakis, P. T. Mathiopoulos, and D. P. Bouras, "Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN–A combined envelope, multiple differential and coherent detection approach," *IEEE Trans. Commun.*, pp. 63–75, January 1994.

[54] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels," *IEEE Trans. Commun.*, pp. 788–99, July 1997.

[55] H. Kong and E. Shwedyk, "Sequence detection and channel state estimation over finite state Markov channels," *IEEE Trans. Veh. Tech.*, pp. 833–9, May 1999.

[56] G. M. Vitetta and D. P. Taylor, "Maximum-likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, pp. 2750–8, November 1995.

[57] L. Li and A. J. Goldsmith, "Low-complexity maximum-likelihood detection of coded signals sent over finite-state Markov channels," *IEEE Trans. Commun.*, pp. 524–31, April 2002.

[58] E. K. Hall and S. G. Wilson, "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, pp. 160–74, February 1998.

[59] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Sel. Areas Commun*. pp. 1706–17, September 2001.

[60] J. Hou, P. H. Siegel, and L. B. Milstein, "Performance analysis and code optimization of low density parity-check codes on Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, pp. 924–34, May 2001.

[61] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inform. Theory*, pp. 868–86, May 1996.

[62] G. D. Forney, "Burst error correcting codes for the classic bursty channel," *IEEE Trans. Commun. Tech.*, pp. 772–81, October 1971.

[63] J. L. Ramsey, "Realization of optimum interleavers," *IEEE Trans. Inform. Theory*, pp. 338–45, 1970.

[64] C.-E. W. Sundberg and N. Seshadri, "Coded modulation for fading channels-An overview," *Euro. Trans. Telecommun.*, pp. 309–24, May/ June 1993.

[65] L.-F. Wei, "Coded M-DPSK with built-in time diversity for fading channels," *IEEE Trans. Inform. Theory*, pp. 1820–39, November 1993.

[66] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*, Kluwer, New York, 1994.

[67] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, pp. 927–46, May 1998.

[68] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, pp. 873–84, May 1992.

[69] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inform. Theory*, pp. 371–7, May 1977.

[70] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "Variable rate sub-band speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Trans. Signal Proc.*, pp. 1717–31, August 1991.

[71] J. W. Modestino and D. G. Daut, "Combined source-channel coding of images," *IEEE Trans. Commun.*, pp. 1644–59, November 1979.

[72] A. R. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," *IEEE Trans. Inform. Theory*, pp. 1234–48, July 1993.

[73] L.-F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, pp. 1439–49, October 1993.

[74] N. Seshadri and C.-E. W. Sundberg, "Multilevel trellis coded modulations for the Rayleigh fading channel," *IEEE Trans. Commun.*, pp. 1300–10, September 1993.

[75] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE National Convention Record*, part 4, pp. 142–63, 1959.

[76] N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," *IEEE J. Sel. Areas Commun.*, pp. 926–43, June 1992.

[77] H. Jafarkhani, P. Ligdas, and N. Farvardin, "Adaptive rate allocation in ajoint source/channel coding framework for wireless channels," *Proc. IEEE Veh. Tech. Conf.*, pp. 492–6, April 1996.

[78] W. C. Wong, R. Steele, and C.-E. W. Sundberg, *Source-Matched Mobile Communications*, Pentech and IEEE Press, London and New York, 1995.

[79] K.-P. Ho and J. M. Kahn, "Transmission of analog signals using multicarrier modulation: A combined source-channel coding approach," *IEEE Trans. Commun.*, pp. 1432–43, November 1996.

[80] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers," *IEEE Trans. Inform. Theory*, pp. 155–60, January 1991.

[81] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source-channel coding," *IEEE Trans. Inform. Theory*, pp. 827–38, November 1987.

[82] K.-P. Ho and J. M. Kahn, "Combined source-channel coding using channel-optimized quantizer and multicanier modulation," *Proc. IEEE Internat. Conf. Commun.*, pp. 1323–7, June 1996.

[83] A. J. Goldsmith and M. Effros, "Joint design of fixed-rate source codes and multiresolution channel codes," *IEEE Trans. Commun.*, pp. 1301–12, October 1998.

[84] E. Ayanoglu and R. M. Gray, "The design of joint source and channel trellis waveform coders," *IEEE Trans. Inform. Theory*, pp. 855–65, November 1987.

[85] T. R. Fischer and M. W. Marcellin, "Joint trellis coded quantization/modulation," *IEEE Trans. Commun.*, pp. 172–6, Febmary 1991.

Chapter 9

Adaptive Modulation and Coding

Adaptive modulation and coding enable robust and spectrally efficient transmission over time-varying channels. The basic premise is to estimate the channel at the receiver and feed this estimate back to the transmitter, so that the transmission scheme can be adapted relative to the channel characteristics. Modulation and coding techniques that do not adapt to fading conditions require a fixed link margin to maintain acceptable performance when the channel quality is poor. Thus, these systems are effectively designed for worst-case channel conditions. Since Rayleigh fading can cause a signal power loss of up to 30 dB, designing for the worst-case channel conditions can result in very inefficient utilization of the channel. Adapting to the channel fading can increase average throughput, reduce required transmit power, or reduce average probability of bit error by taking advantage of favorable channel conditions to send at higher data rates or lower power – and by reducing the data rate or increasing power as the channel degrades. In Section 4.2.4 we derived the optimal adaptive transmission scheme that achieves the Shannon capacity of a flat fading channel. In this chapter we describe more practical adaptive modulation and coding techniques to maximize average spectral efficiency while maintaining a given average or instantaneous bit error probability. The same basic premise can be applied to MIMO channels, frequency-selective fading channels with equalization, OFDM or CDMA, and cellular systems. The application of adaptive techniques to these systems will be described in subsequent chapters.

Adaptive transmission was first investigated in the late sixties and early seventies [1, 2]. Interest in these techniques was short-lived, perhaps due to hardware constraints, lack of good channel estimation techniques, and or systems focusing on point-to-point radio links without transmitter feedback. As technology evolved these issues became less constraining, resulting in a revived interest in the theory and practical application of adaptive transmission methods starting in the 1990s [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. The most prevalent wireless systems today, including cellular and Wi-Fi networks, use the adaptive MQAM techniques of [12] coupled with adaptive coding, as described in more detail in Appendix D.

There are several practical constraints that determine when adaptive modulation should be used. Adaptive modulation requires a feedback path between the transmitter and receiver, which is typical in most but not all wireless system designs. If the channel is changing faster than it can be reliably estimated and fed back to the transmitter, adaptive techniques will perform poorly. Many wireless channels exhibit variations on different time scales; examples include multipath fading, which can change very quickly, and shadowing, which changes more slowly. Often only the slow variations can be tracked and adapted to, in which case flat fading mitigation is needed to address the effects of multipath. Hardware constraints may dictate how often the transmitter can change its rate and/or power, and this may limit the performance gains possible with adaptive modulation. Finally, adaptive modulation typically varies the rate of data transmission relative to channel conditions. We will see that average spectral efficiency of adaptive modulation under an average power constraint is maximized by setting the data rate to be small or zero in poor channel conditions. However, with this scheme the quality of fixed-rate applications

(such as voice or video) with hard delay constraints may be significantly compromised. Thus, in delay-constrained applications the adaptive modulation should be optimized to minimize outage probability for a fixed data rate [13].

9.1 Adaptive Transmission System

In this section we describe the system associated with adaptive transmission. The model is the same as the narrowband fading model of Section 4.2.1 used to determine the capacity of flat fading channels. We assume linear modulation, where the adaptation takes place at a multiple of the symbol rate $R_s = 1/T_s$. We also assume the modulation uses ideal Nyquist data pulses (sinc$[t/T_s]$), so the signal bandwidth $B = 1/T_s$. We model the flat fading channel as a discrete-time channel in which each channel use corresponds to one symbol time T_s. The channel has stationary and ergodic time-varying gain $\sqrt{g[i]}$ that follows a given distribution $p(g)$ and AWGN $n[i]$, with power spectral density $N_0/2$. Let $\bar{P}$ denote the average transmit signal power, $B = 1/T_s$ the received signal bandwidth, and $\bar{g}$ the average channel gain. The instantaneous received signal-to-noise power ratio (SNR) is then $\gamma[i] = \bar{P}g[i]/N_0B$, $0 \leq \gamma[i] < \infty$, and its expected value over all time is $\bar{\gamma} = \bar{P}\bar{g}/N_0B$. Since $g[i]$ is stationary, the distribution of $\gamma[i]$ is independent of i, and we denote this distribution by $p(\gamma)$.

In adaptive transmission we estimate the power gain or the received SNR at time i and then adapt the modulation and coding parameters accordingly. The most common parameters to adapt are the data rate $R[i]$, transmit power $P[i]$, and coding parameters $C[i]$. For M-ary modulation the data rate $R[i] = \log_2 M[i]/T_s = B\log_2 M[i]$ bps. The *spectral efficiency* of the M-ary modulation is $R[i]/B = \log_2 M[i]$ bps/Hz. We denote the SNR estimate as $\hat{\gamma}[i] = \bar{P}\hat{g}[i]/N_0B$, which is based on the power gain estimate $\hat{g}[i]$. Suppose the transmit power is adapted relative to $\hat{\gamma}[i]$. We denote this adaptive transmit power at time i by $P(\hat{\gamma}[i]) = P[i]$, and the received power at time i is then $\gamma[i]P(\hat{\gamma}[i])/\bar{P}$. Similarly, we can adapt the data rate of the modulation $R(\hat{\gamma}[i]) = R[i]$ and/or the coding parameters $C(\hat{\gamma}[i]) = C[i]$ relative to the estimate $\hat{\gamma}[i]$. When the context is clear, we will omit the time reference i, including in the terms γ, $P(\gamma)$, $R(\gamma)$, and $C(\gamma)$.

The system model is illustrated in Figure 9.1. We assume that an estimate $\hat{g}[i]$ of the channel power gain $g[i]$ at time i is available to the receiver after an estimation time delay of i_e and that this same estimate is available to the transmitter after a combined estimation and feedback path delay of $i_d = i_e + i_f$. The availability of this channel information at the transmitter allows it to adapt its transmission scheme relative to the channel variation. The adaptive strategy may take into account the estimation error and delay in $\hat{g}[i - i_d]$ or it may treat $\hat{g}[i - i_d]$ as the true gain: this issue will be discussed in more detail in Section 9.3.7. We assume that the feedback path does not introduce any errors, which is a reasonable assumption if strong error correction and detection codes are used on the feedback path and if packets associated with detected errors are retransmitted.

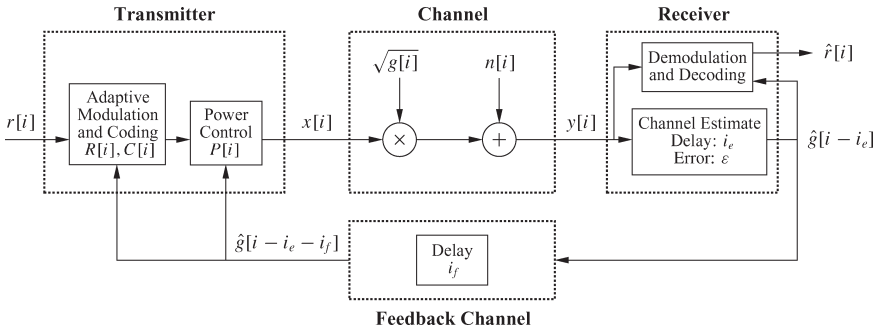


Figure 9.1: System model.

The rate of channel variation will dictate how often the transmitter must adapt its transmission parameters, and it will also affect the estimation error of $g[i]$. When the channel gain consists of both fast and slow fading

components, the adaptive transmission may adapt to both if $g[i]$ changes sufficiently slowly, or it may adapt to just the slow fading. In particular, if $g[i]$ corresponds to shadowing and multipath fading, then at low speeds the shadowing is essentially constant, and the multipath fading is sufficiently slow that it can be estimated and fed back to the transmitter with estimation error and delay that do not significantly degrade performance. At high speeds the system can no longer effectively estimate and feed back the multipath fading in order to adapt to it. In this case, the adaptive transmission responds to the shadowing variations only, and the error probability of the modulation must be averaged over the fast fading distribution. Adaptive techniques for combined fast and slow fading are discussed in Section 9.5.

9.2 Adaptive Techniques

There are many parameters that can be varied at the transmitter relative to the channel gain γ. In this section we discuss adaptive techniques associated with variation of the most common parameters: data rate, power, coding, error probability, and combinations of these adaptive techniques.

9.2.1 Variable-Rate Techniques

In variable-rate modulation the data rate $R[\gamma]$ is varied relative to the channel gain γ. This can be done by fixing the symbol rate $R_s = 1/T_s$ of the modulation and using multiple modulation schemes or constellation sizes, or by fixing the modulation (e.g. BPSK) and changing the symbol rate. Symbol rate variation is difficult to implement in practice because a varying signal bandwidth is impractical and complicates bandwidth sharing. In contrast, changing the constellation size or modulation type with a fixed symbol rate is fairly easy, and these techniques are used in current systems. Specifically, the GSM and IS-136 EDGE system as well as 802.11a wireless LANs vary their modulation and coding relative to channel quality [**?**]. In general the modulation parameters that dictate the transmission rate are fixed over a block or frame of symbols, where the frame size is a parameter of the design. Frames may also include pilot symbols for channel estimation and other control information.

When a discrete set of modulation types or constellation sizes are used, each value of γ must be mapped to one of the possible modulation schemes. This is often done to maintain the bit error probability of each scheme below a given value. These ideas are illustrated in the following example as well as in subsequent sections on specific adaptive modulation techniques.

Example 9.1: Consider an adaptive modulation system that uses QPSK and 8-PSK for a target P_b of approximately 10^{-3}. If the target P_b cannot be met with either scheme, then no data is transmitted. Find the range of γ-values associated with the three possible transmission schemes (no transmission, QPSK, and 8-PSK) as well as the average spectral efficiency of the system, assuming Rayleigh fading with $\bar{\gamma} = 20$ dB.

Solution: First note that the SNR $\gamma = \gamma_s$ for both QPSK and 8-PSK. From Section 6.1 we have $P_b \approx Q(\sqrt{\gamma})$ for QPSK and $P_b \approx .666Q(\sqrt{2\gamma}\sin(\pi/8))$ for 8-PSK. Since $\gamma > 14.79$ dB yields $P_b < 10^{-3}$ for 8-PSK, the adaptive modulation uses 8-PSK modulation for $\gamma > 14.79$ dB. Since $\gamma > 10.35$ dB yields $P_b < 10^{-3}$ for QPSK, the adaptive modulation uses QPSK modulation for 14.79 dB $\geq \gamma > 10.35$ dB. The channel is not used for $\gamma \leq 10.35$ dB.

We determine the average rate by analyzing how often each of the different transmission schemes is used. Since 8-PSK is used when $\gamma > 14.79$ dB = 30.1, in Rayleigh fading with $\bar{\gamma} = 20$ dB the spectral efficiency $R[\gamma]/B = \log_2 8 = 3$ bps/Hz is transmitted a fraction of time equal to $P_8 = \int_{30.1}^{\infty} \frac{1}{100} e^{-\gamma/100}\, d\gamma = .74$. QPSK is used when $10.35 < \gamma \leq 14.79$ dB, where 10.35 dB = 10.85 in linear units. So $R[\gamma] = \log_2 4 = 2$ bps/Hz is transmitted a fraction of time equal to $P_4 = \int_{10.85}^{30.1} \frac{1}{100} e^{-\gamma/100}\, d\gamma = .157$. During the remaining .103 portion of time there is no data transmission. So the average spectral efficiency is $.74 \cdot 3 + .157 \cdot 2 + .103 \cdot 0 = 2.534$ bps/Hz.

Note that if $\gamma \leq 10.35$ dB then – rather than suspending transmission, which leads to an outage probability of roughly .1 – either just one signaling dimension could be used (i.e., BPSK could be transmitted) or error correction coding could be added to the QPSK to meet the P_b target. If block or convolutional codes were used then the spectral efficiency for $\gamma \leq 10.35$ dB would be less than 2 bps/Hz but larger than a spectral efficiency of zero corresponding to no transmission. These variable-coding techniques are described in Section 9.2.4.

9.2.2 Variable-Power Techniques

Adapting the transmit power alone is generally used to compensate for SNR variation due to fading. The goal is to maintain a fixed bit error probability or, equivalently, a constant received SNR. The power adaptation thus inverts the channel fading so that the channel appears as an AWGN channel to the modulator and demodulator.[1] The power adaptation for channel inversion is given by

$$\frac{P(\gamma)}{\bar{P}} = \frac{\sigma}{\gamma}, \tag{9.1}$$

where σ equals the constant received SNR. The average power constraint $\bar{P}$ implies that

$$\int \frac{P(\gamma)}{\bar{P}} p(\gamma)\, d\gamma = \int \frac{\sigma}{\gamma} p(\gamma)\, d\gamma = 1. \tag{9.2}$$

Solving (9.2) for σ yields that $\sigma = 1/\mathbf{E}[1/\gamma]$, so σ is determined by $p(\gamma)$, which in turn depends on the average transmit power $\bar{P}$ through $\bar{\gamma}$. Thus, for a given average power $\bar{P}$, if the value for σ required to meet the target bit error rate is greater than $1/\mathbf{E}[1/\gamma]$ then this target cannot be met. Note that for Rayleigh fading, where γ is exponentially distributed, $\mathbf{E}[1/\gamma] = \infty$ and so no target P_b can be met using channel inversion.

The fading can also be inverted above a given cutoff γ_0, which leads to a truncated channel inversion for power adaptation. In this case the power adaptation is given by

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \sigma/\gamma & \gamma \geq \gamma_0, \\ 0 & \gamma < \gamma_0. \end{cases} \tag{9.3}$$

The cutoff value γ_0 can be based on a desired outage probability $P_{\text{out}} = p(\gamma < \gamma_0)$ or on a desired target BER above a cutoff that is determined by the target BER and $p(\gamma)$. Since the channel is only used when $\gamma \geq \gamma_0$, given an average power $\bar{P}$ we have $\sigma = 1/\mathbf{E}_{\gamma_0}[1/\gamma]$, where

$$\mathbf{E}_{\gamma_0}\left[\frac{1}{\gamma}\right] \triangleq \int_{\gamma_0}^{\infty} \frac{1}{\gamma} p(\gamma)\, d\gamma. \tag{9.4}$$

Example 9.2: Find the power adaptation for BPSK modulation that maintains a fixed $P_b = 10^{-3}$ in non-outage for a Rayleigh fading channel with $\bar{\gamma} = 10$ dB. Also find the resulting outage probability.
Solution: The power adaptation is truncated channel inversion, so we need only find σ and γ_0. For BPSK modulation, with a constant SNR of $\sigma = 4.77$ we get $P_b = Q(\sqrt{2\sigma}) = 10^{-3}$. Setting $\sigma = 1/\mathbf{E}_{\gamma_0}[1/\gamma]$ and solving for γ_0, which must be done numerically, yields $\gamma_0 = .7423$. So $P_{\text{out}} = p(\gamma < \gamma_0) = 1 - e^{-\gamma_0/10} = .379$. Hence there is a high outage probability, which results from requiring $P_b = 10^{-3}$ in this relatively weak channel.

---

[1] Channel inversion and truncated channel inversion were discussed in Section 4.2.4 in the context of fading channel capacity.

9.2.3 Variable Error Probability

We can also adapt the instantaneous BER subject to an average BER constraint $\bar{P}_b$. In Section 6.3.2 we saw that in fading channels the instantaneous error probability varies as the received SNR γ varies, resulting in an average BER of $\bar{P}_b = \int P_b(\gamma)p(\gamma)\,d\gamma$. This is not considered an adaptive technique, since the transmitter does not adapt to γ. Thus, in adaptive modulation, error probability is typically adapted along with some other form of adaptation such as constellation size or modulation type. Adaptation based on varying both data rate and error probability to reduce transmit energy was first proposed by Hayes in [1], where a 4-dB power savings was obtained at a target average bit error probability of 10^{-4}.

9.2.4 Variable-Coding Techniques

In adaptive coding, different channel codes are used to provide different amounts of coding gain to the transmitted bits. For example, a stronger error correction code may be used when γ is small, with a weaker code or no coding used when γ is large. Adaptive coding can be implemented by multiplexing together codes with different error correction capabilities. However, this approach requires that the channel remain roughly constant over the block length or constraint length of the code [7]. On such slowly varying channels, adaptive coding is particularly useful when the modulation must remain fixed, as may be the case owing to complexity or peak-to-average power ratio constraints.

An alternative technique to code multiplexing is rate-compatible punctured convolutional (RCPC) codes [14]. This is a family of convolutional codes at different code rates $R_c = k/n$. The basic premise of RCPC codes is to have a single encoder and decoder whose error correction capability can be modified by not transmitting certain coded bits (*puncturing* the code). Moreover, RCPC codes have a rate compatibility constraint so that the coded bits associated with a high-rate (weaker) code are also used by all lower-rate (stronger) codes. Thus, to increase the error correction capability of the code, the coded bits of the weakest code are transmitted along with additional coded bits to achieve the desired level of error correction. The rate compatibility makes it easy to adapt the error protection of the code, since the same encoder and decoder are used for all codes in the RCPC family, with puncturing at the transmitter to achieve the desired error correction. Decoding is performed by a Viterbi algorithm operating on the trellis associated with the lowest rate code, with the puncturing incorporated into the branch metrics. Puncturing is a very effective and powerful adaptive coding technique, and it forms the basis of adaptive coding in the GSM and IS-136 EDGE protocol for data transmission [**?**].

Adaptive coding through either multiplexing or puncturing can be done for fixed modulation or combined with adaptive modulation as a hybrid technique. When the modulation is fixed, adaptive coding is often the only practical mechanism to address the channel variations [6, 7]. The focus of this chapter is on systems where adaptive modulation is possible, so adaptive coding on its own will not be further discussed.

9.2.5 Hybrid Techniques

Hybrid techniques can adapt multiple parameters of the transmission scheme, including rate, power, coding, and instantaneous error probability. In this case joint optimization of the different techniques is used to meet a given performance requirement. Rate adaptation is often combined with power adaptation to maximize spectral efficiency, and we apply this joint optimization to different modulations in subsequent sections.

9.3 Variable-Rate Variable-Power MQAM

In the previous section we discussed general approaches to adaptive modulation and coding. In this section we describe a specific form of adaptive modulation where the rate and power of MQAM are varied to maximize spec-

tral efficiency while meeting a given instantaneous P_b target. We study this specific form of adaptive modulation because it provides insight into the benefits of adaptive modulation and since, moreover, the same scheme for power and rate adaptation that achieves capacity also optimizes this adaptive MQAM design. We will show that there is a constant power gap between the spectral efficiency of this adaptive MQAM technique and capacity in flat fading, and this gap can be partially closed by coded modulation using a trellis or lattice code superimposed on the adaptive modulation.

Consider a family of MQAM signal constellations with a fixed symbol time T_s, where M denotes the number of points in each signal constellation. We assume $T_s = 1/B$ based on ideal Nyquist pulse shaping. Let $\bar{P}$, N_0, $\gamma = \bar{P}g/N_0B$, and $\bar{\gamma} = \bar{P}\bar{g}/N_0B$ be as given in our system model. Then the average E_s/N_0 equals the average SNR:

$$\frac{\bar{E}_s}{N_0} = \frac{\bar{P}\bar{g}T_s}{N_0} = \bar{\gamma}. \tag{9.5}$$

The spectral efficiency for fixed M is $R/B = \log_2 M$, the number of bits per symbol. This efficiency is typically parameterized by the average transmit power $\bar{P}$ and the BER of the modulation technique.

9.3.1 Error Probability Bounds

In [15] the BER for an AWGN channel with MQAM modulation, ideal coherent phase detection, and SNR γ is bounded by

$$P_b \leq 2e^{-1.5\gamma/(M-1)}. \tag{9.6}$$

A tighter bound, good to within 1 dB for $M \geq 4$ and $0 \leq \gamma \leq 30$ dB, is

$$P_b \leq .2e^{-1.5\gamma/(M-1)}. \tag{9.7}$$

Note that these expressions are only bounds, so they don't match the error probability expressions from Table 6.1. We use these bounds because they are easy to invert: so we can obtain M as a function of the target P_b and the power adaptation policy, as we will see shortly. Adaptive modulation designs can also be based on BER expressions that are not invertible or on BER simulation results, with numerical inversion used to obtain the constellation size and SNR associated with a given BER target.

In a fading channel with nonadaptive transmission (constant transmit power and rate), the average BER is obtained by integrating the BER in AWGN over the fading distribution $p(\gamma)$. Thus, we use the average BER expression to find the maximum data rate that can achieve a given average BER for a given average SNR. Similarly, if the data rate and average BER are fixed, we can find the required average SNR to achieve this target, as illustrated in the next example.

Example 9.3: Find the average SNR required to achieve an average BER of $\bar{P}_b = 10^{-3}$ for nonadaptive BPSK modulation in Rayleigh fading. What is the spectral efficiency of this scheme?
Solution: From Section 6.3.2, BPSK in Rayleigh fading has $\bar{P}_b \approx 1/4\bar{\gamma}$. Thus, without transmitter adaptation, for a target average BER of $\bar{P}_b = 10^{-3}$ we require $\bar{\gamma} = 1/4\bar{P}_b = 250 = 24$ dB. The spectral efficiency is $R/B = \log_2 2 = 1$ bps/Hz. We will see that adaptive modulation provides a much higher spectral efficiency at this same SNR and target BER.

9.3.2 Adaptive Rate and Power Schemes

We now consider adapting the transmit power $P(\gamma)$ relative to γ, subject to the average power constraint $\bar{P}$ and an instantaneous BER constraint $P_b(\gamma) = P_b$. The received SNR is then $\gamma P(\gamma)/\bar{P}$, and the P_b bound for each value of γ, using the tight bound (9.7), becomes

$$P_b(\gamma) \leq .2 \exp\left[\frac{-1.5\gamma}{M-1}\frac{P(\gamma)}{\bar{P}}\right]. \tag{9.8}$$

We adjust M and $P(\gamma)$ to maintain the target P_b. Rearranging (9.8) yields the following maximum constellation size for a given P_b:

$$M(\gamma) = 1 + \frac{1.5\gamma}{-\ln(5P_b)}\frac{P(\gamma)}{\bar{P}} = 1 + K\gamma\frac{P(\gamma)}{\bar{P}}, \tag{9.9}$$

where

$$K = \frac{-1.5}{\ln(5P_b)} < 1. \tag{9.10}$$

We maximize spectral efficiency by maximizing

$$\mathbf{E}[\log_2 M(\gamma)] = \int_0^\infty \log_2\left(1 + \frac{K\gamma P(\gamma)}{\bar{P}}\right) p(\gamma)\, d\gamma \tag{9.11}$$

subject to the power constraint

$$\int_0^\infty P(\gamma)p(\gamma)\, d\gamma = \bar{P}. \tag{9.12}$$

The power adaptation policy that maximizes (9.11) has the same form as the optimal power adaptation policy (4.16) that achieves capacity:

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} 1/\gamma_0 - 1/(\gamma K) & \gamma \geq \gamma_0/K, \\ 0 & \gamma < \gamma_0/K, \end{cases} \tag{9.13}$$

where γ_0/K is the optimized cutoff fade depth below which the channel is not used, for K given by (9.10). If we define $\gamma_K = \gamma_0/K$ and multiply both sides of (9.13) by K, we obtain

$$\frac{KP(\gamma)}{\bar{P}} = \begin{cases} 1/\gamma_K - 1/\gamma & \gamma \geq \gamma_K, \\ 0 & \gamma < \gamma_K, \end{cases} \tag{9.14}$$

where γ_K is a cutoff fade depth below which the channel is not used. This cutoff must satisfy the power constraint (9.12):

$$\int_{\gamma_K}^\infty \left(\frac{1}{\gamma_K} - \frac{1}{\gamma}\right) p(\gamma)\, d\gamma = K. \tag{9.15}$$

Substituting (9.13) or (9.14) into (9.9), we get that the instantaneous rate is given by

$$\log_2 M(\gamma) = \log_2\left(\frac{\gamma}{\gamma_K}\right), \tag{9.16}$$

and the corresponding average spectral efficiency (9.11) is given by

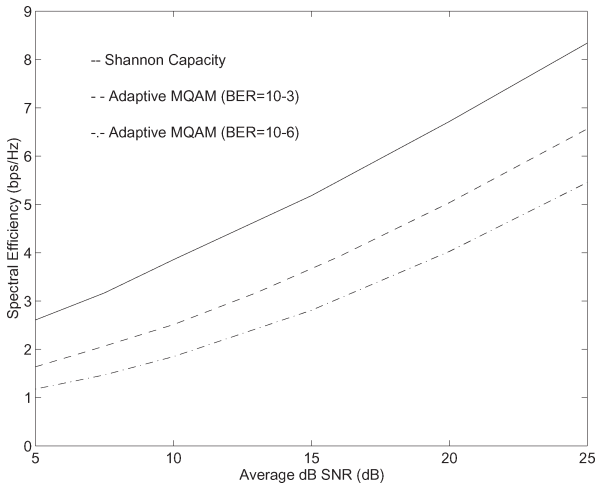


Figure 9.2: Average spectral efficiency in log-normal shadowing ($\sigma_{\psi_{\rm dB}} = 8$ dB).

$$\frac{R}{B} = \int_{\gamma_K}^{\infty} \log_2 \left(\frac{\gamma}{\gamma_K} \right) p(\gamma)\, d\gamma. \tag{9.17}$$

Comparing the power adaptation and average spectral efficiency – (4.16) and (4.17) – associated with the Shannon capacity of a fading channel with (9.13) and (9.17), the optimal power adaptation and average spectral efficiency of adaptive MQAM, we see that the power and rate adaptation are the same and lead to the same average spectral efficiency, with an effective power loss of K for adaptive MQAM as compared to the capacity-achieving scheme. Moreover, this power loss is independent of the fading distribution. Thus, if the capacity of a fading channel is R bps/Hz at SNR $\bar{\gamma}$, then uncoded adaptive MQAM requires a received SNR of $\bar{\gamma}/K$ to achieve the same rate. Equivalently, K is the maximum possible coding gain for this variable rate and power MQAM method. We discuss superimposing a trellis or lattice code on the adaptive MQAM to obtain some of this coding gain in Section 9.3.8.

We plot the average spectral efficiency (9.17) of adaptive MQAM at a target P_b of 10^{-3} and 10^{-6} for both log-normal shadowing and Rayleigh fading in Figures 9.2 and 9.3, respectively. We also plot the capacity in these figures for comparison. Note that the gap between the spectral efficiency of variable-rate variable-power MQAM and capacity is the constant K, which from (9.10) is a simple function of the BER.

9.3.3 Channel Inversion with Fixed Rate

We can also apply channel inversion power adaptation to maintain a fixed received SNR. We then transmit a single fixed-rate MQAM modulation that achieves the target P_b. The constellation size M that meets this target P_b is obtained by substituting the channel inversion power adaptation $P(\gamma)/\bar{P} = \sigma/\gamma$ of (9.2) into (9.9) with $\sigma = 1/\mathbf{E}[1/\gamma]$. Since the resulting spectral efficiency $R/B = M$, this yields the spectral efficiency of the channel inversion power adaptation as

$$\frac{R}{B} = \log_2 \left(1 + \frac{-1.5}{\ln(5P_b)\mathbf{E}[1/\gamma]} \right). \tag{9.18}$$

This spectral efficiency is based on the tight bound (9.7); if the resulting $M = R/B < 4$ then the loose bound (9.6) must be used, in which case $\ln(5P_b)$ is replaced by $\ln(.5P_b)$ in (9.18).

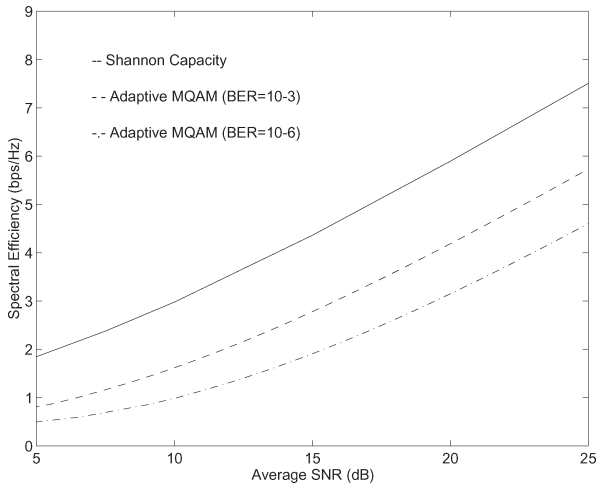


Figure 9.3: Average spectral efficiency in Rayleigh fading.

With truncated channel inversion the channel is only used when $\gamma > \gamma_0$. Thus, the spectral efficiency with truncated channel inversion is obtained by substituting $P(\gamma)/\bar{P} = \sigma/\gamma$, $\gamma > \gamma_0$, into (9.9) and multiplying by the probability that $\gamma > \gamma_0$. The maximum value is obtained by optimizing relative to the cutoff level γ_0:

$$\frac{R}{B} = \max_{\gamma_0} \log_2 \left(1 + \frac{-1.5}{\ln(5P_b)\mathbf{E}_{\gamma_0}[1/\gamma]} \right) p(\gamma > \gamma_0). \tag{9.19}$$

The spectral efficiency of adaptive MQAM (with the optimal water-filling and truncated channel inversion power adaptation) in a Rayleigh fading channel with a target BER of 10^{-3} is shown in Figure 9.4, along with the capacity under the same two power adaptation policies. We see, surprisingly, that truncated channel inversion with fixed-rate transmission has almost the same spectral efficiency as optimal variable-rate variable-power MQAM. This suggests that truncated channel inversion is more desirable in practice, since it achieves almost the same spectral efficiency as variable-rate variable-power transmission but does not require that we vary the rate. However, this assumes there is no restriction on constellation size. Specifically, the spectral efficiencies (9.17), (9.18), and (9.19) assume that M can be any real number and that the power and rate can vary continuously with γ. Though MQAM modulation for noninteger values of M is possible, the complexity is quite high [16]. Moreover, it is difficult in practice to continually adapt the transmit power and constellation size to the channel fading, particularly in fast fading environments. Thus, we now consider restricting the constellation size to just a handful of values. This will clearly affect the spectral efficiency, though (as shown in the next section) not by very much.

9.3.4 Discrete-Rate Adaptation

We now assume the same model as in the previous section, but we restrict the adaptive MQAM to a limited set of constellations, as is necessary for practical implementation. Specifically, we assume a set of square constellations of size $M_0 = 0, M_1 = 2$, and $M_j = 2^{2(j-1)}$, $j = 2, \ldots, N-1$, for some N. We assume square constellations for $M > 2$ since they are easier to implement than rectangular ones [17]. We first analyze the impact of this restriction on the spectral efficiency of the optimal adaptation policy. We then determine the effect on the channel inversion policies.

Consider a variable-rate variable-power MQAM transmission scheme subject to these constellation restrictions. Thus, at each symbol time we transmit a symbol from a constellation in the set $\{M_j \colon j = 0, 1, \ldots, N-1\}$; the choice of constellation depends on the fade level γ over that symbol time. Choosing the M_0 constellation

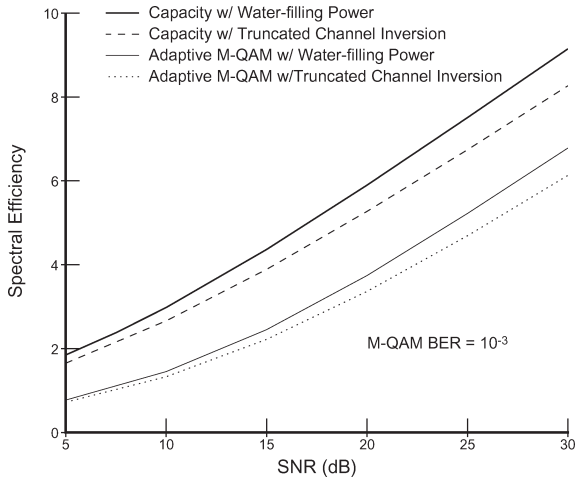


Figure 9.4: Spectral efficiency with different power adaptation policies (Rayleigh fading).

corresponds to no data transmission. For each value of γ, we must decide which constellation to transmit and what the associated transmit power should be. The rate at which the transmitter must change its constellation and power is analyzed below. Since the power adaptation is continuous while the constellation size is discrete, we call this a continuous-power discrete-rate adaptation scheme.

We determine the constellation size associated with each γ by discretizing the range of channel fade levels. Specifically, we divide the range of γ into *N fading regions* $R_j = [\gamma_{j-1}, \gamma_j)$, $j = 0, \ldots, N-1$, where $\gamma_{-1} = 0$ and $\gamma_{N-1} = \infty$. We transmit constellation M_j when $\gamma \in R_j$. The spectral efficiency for $\gamma \in R_j$ is thus $\log_2 M_j$ bps/Hz for $j > 0$.

The adaptive MQAM design requires that the boundaries of the R_j regions be determined. Although these boundaries can be optimized to maximize spectral efficiency, as derived in Section 9.4.2, the optimal boundaries cannot be found in closed form and require an exhaustive search to obtain. Thus, we will use a suboptimal technique to determine boundaries. These suboptimal boundaries are much easier to find than the optimal ones and have almost the same performance. Define

$$M(\gamma) = \frac{\gamma}{\gamma_K^*}, \tag{9.20}$$

where $\gamma_K^* > 0$ is a parameter that will later be optimized to maximize spectral efficiency. Note that substituting (9.13) into (9.9) yields (9.20) with $\gamma_K^* = \gamma_K$. Therefore, the appropriate choice of γ_K^* in (9.20) defines the optimal constellation size for each γ when there is no constellation restriction.

Assume now that γ_K^* is fixed and define $M_N = \infty$. To obtain the constellation size M_j, $j = 0, \ldots, N-1$, for a given SNR γ, we first compute $M(\gamma)$ from (9.20). We then find j such that $M_j \leq M(\gamma) < M_{j+1}$ and assign constellation M_j to this γ-value. Thus, for a fixed γ, we transmit the largest constellation in our set $\{M_j \colon j = 0, \ldots, N-1\}$ that is smaller than $M(\gamma)$. For example, if the fade level γ satisfies $2 \leq \gamma/\gamma_K^* < 4$ then we transmit BPSK. The region boundaries other than $\gamma_{-1} = 0$ and $\gamma_{N-1} = \infty$ are located at $\gamma_j = \gamma_K^* M_{j+1}$, $j = 0, \ldots, N-2$. This mapping of γ values to fading regions R_j and associated constellations M_j is shown in Figure 9.5. Clearly, increasing the number of discrete signal constellations N yields a better approximation to the continuous adaptation (9.9), resulting in a higher spectral efficiency.

Once the regions and associated constellations are fixed, we must find a power adaptation policy that satisfies the BER requirement and the power constraint. By (9.9) we can maintain a fixed BER for the constellation $M_j > 0$ using the power adaptation policy

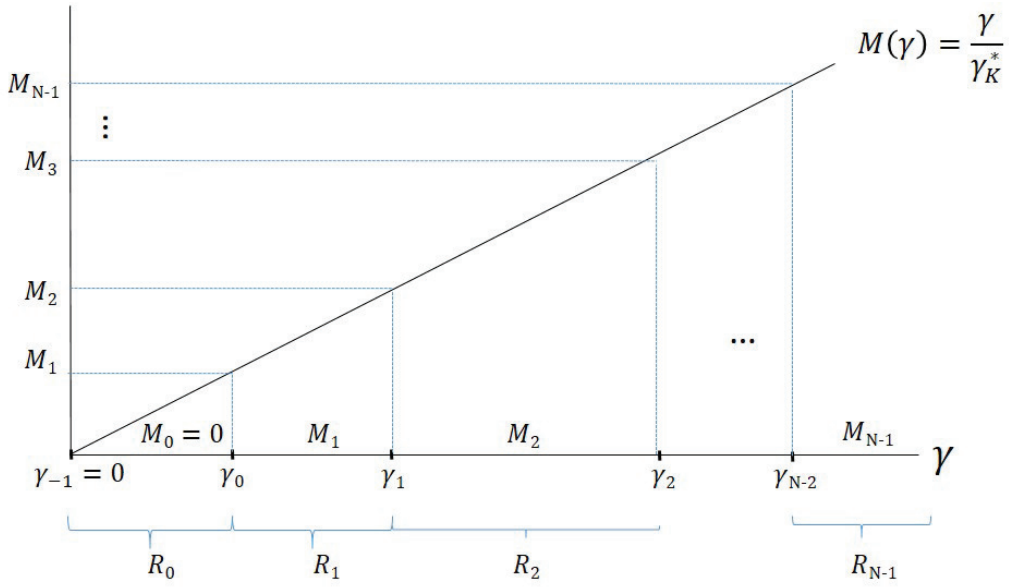


Figure 9.5: Mapping γ to Fading Regions R_j and Associated Constellations M_j in Discrete-Rate Adaptation

$$\frac{P_j(\gamma)}{\bar{P}} = \begin{cases} (M_j - 1)/(\gamma K) & M_j \le \gamma/\gamma_K^* < M_{j+1}, \\ 0 & M_j = 0 \end{cases} \tag{9.21}$$

for $\gamma \in R_j$, since this power adaptation policy leads to a fixed received E_s/N_0 for the constellation M_j of

$$\frac{E_s(j)}{N_0} = \frac{\gamma P_j(\gamma)}{\bar{P}} = \frac{M_j - 1}{K}. \tag{9.22}$$

By definition of K, MQAM modulation with constellation size M_j and E_s/N_0 given by (9.22) results in the desired target P_b. In Table 9.1 we tabulate the constellation size and power adaptation as a function of γ and γ_K^* for five fading regions.

Table 9.1: Rate and power adaptation for five regions

Region (j)	**γ range**	M_j	$P_j(\gamma)/\bar{P}$
0	$0 \le \gamma/\gamma_K^\star < 2$	0	0
1	$2 \le \gamma/\gamma_K^\star < 4$	2	$1/(\gamma K)$
2	$4 \le \gamma/\gamma_K^\star < 16$	4	$3/(\gamma K)$
3	$16 \le \gamma/\gamma_K^\star < 64$	16	$15/(\gamma K)$
4	$64 \le \gamma/\gamma_K^\star < \infty$	64	$63/(\gamma K)$

The spectral efficiency for this discrete-rate policy is just the sum of the data rates associated with each of the regions multiplied by the probability that γ falls in that region:

$$\frac{R}{B} = \sum_{j=1}^{N-1} \log_2(M_j) p\left(M_j \le \frac{\gamma}{\gamma_K^*} < M_{j+1}\right). \tag{9.23}$$

Since M_j is a function of γ_K^*, we can maximize (9.23) relative to γ_K^*, subject to the power constraint

$$\sum_{j=1}^{N-1} \int_{\gamma_K^* M_j}^{\gamma_K^* M_{j+1}} \frac{P_j(\gamma)}{\bar{P}} p(\gamma)\, d\gamma = 1, \tag{9.24}$$

where $P_j(\gamma)/\bar{P}$ is defined in (9.21). There is no closed-form solution for the optimal γ_K^*: in the calculations that follow it was found using numerical search techniques.

In Figures 9.6 and 9.7 we show the maximum of (9.23) versus the number of fading regions N for log-normal shadowing and Rayleigh fading, respectively. We assume a BER of 10^{-3} for both plots. From Figure 9.6 we see that restricting our adaptive policy to just six fading regions $(M_j = 0, 2, 4, 16, 64, 256)$ results in a spectral efficiency that is within 1 dB of the efficiency obtained with continuous-rate adaptation (9.17) under log-normal shadowing. A similar result holds for Rayleigh fading using five regions $(M_j = 0, 2, 4, 16, 64)$. Essentially this indicates that with 5 regions, $M_j \approx M(\gamma) = \gamma/\gamma_K$, $j = 0, \ldots, N-1$, for $M(\gamma)$ the optimal rate adaptation with no constraints on the constellation size. Both these figures also show that a smaller number of regions leads to a lower average rate. That is because the approximation $M_j \approx M(\gamma) = \gamma/\gamma_K$ no longer holds, both because the smaller number of regions implies that M_j maintains a constant value for a larger range of γ values and because the maximum constellation $M_{N-1} \ll M(\gamma)$ at high SNRs. This latter effect is evident in both figures as, for high SNRs, the average rate levels off to that of the largest constellation in the available set of constellations (e.g. for 4 regions, the maximum constellation is $M_3 = 16$QAM whose spectral efficiency is 4 bps/Hz).

We can simplify our discrete-rate policy even further by using a constant transmit power for each constellation M_j. Thus, each fading region is associated with one signal constellation and one transmit power. This policy is called discrete-power discrete-rate adaptive MQAM. Since the transmit power and constellation size are fixed in each region, the BER will vary with γ in each region. Thus, the region boundaries and transmit power must be set to achieve a given target average BER.

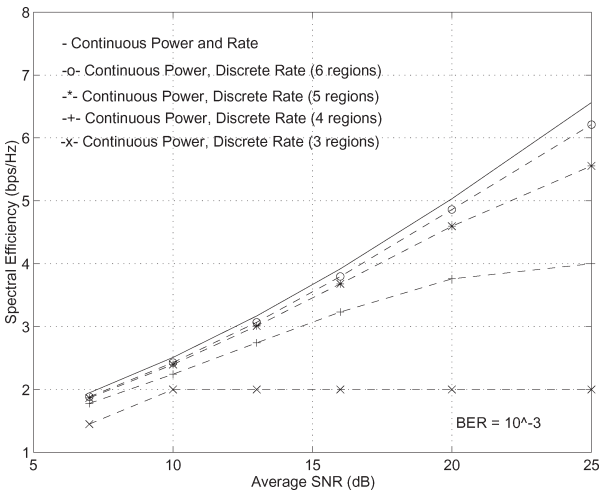


Figure 9.6: Discrete-rate efficiency in log-normal shadowing ($\sigma_{\psi_{\rm dB}} = 8$ dB).

A restriction on allowable signal constellations will also affect the total channel inversion and truncated channel inversion policies. Specifically, suppose that, with the channel inversion policies, the constellation must be chosen from a fixed set of possible constellations $\mathcal{M} = \{M_0 = 0,\ \ldots,\ M_{N-1}\}$. For total channel inversion the spectral efficiency with this restriction is thus

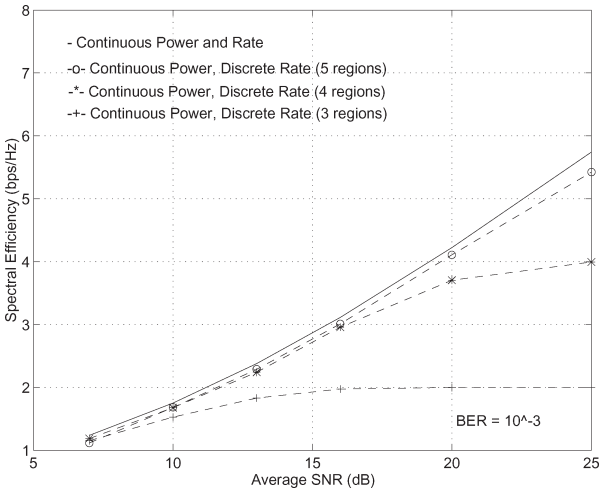


Figure 9.7: Discrete-rate efficiency in Rayleigh fading.

$$\frac{R}{B} = \log_2 \left\lfloor \left(1 + \frac{-1.5}{\ln(5P_b)\mathbf{E}[1/\gamma]}\right)\right\rfloor_{\mathcal{M}}, \tag{9.25}$$

where $\lfloor x \rfloor_{\mathcal{M}}$ denotes the largest number in the set $\mathcal{M}$ less than or equal to x. The spectral efficiency with this policy will be restricted to values of $\log_2 M$, $M \in \mathcal{M}$, with discrete jumps at the $\bar{\gamma}$-values where the spectral efficiency without constellation restriction (9.18) equals $\log_2 M$. For truncated channel inversion the spectral efficiency is given by

$$\frac{R}{B} = \max_{\gamma_0} \log_2 \left\lfloor \left(1 + \frac{-1.5}{\ln(5P_b)\mathbf{E}_{\gamma_0}[1/\gamma]}\right)\right\rfloor_{\mathcal{M}} p(\gamma > \gamma_0). \tag{9.26}$$

In Figures 9.8 and 9.9 we show the impact of constellation restriction on adaptive MQAM for the different power adaptation policies. When the constellation is restricted we assume six fading regions, so $\mathcal{M} = \{M_0 = 0, 2, 4, \ldots, 256\}$. The power associated with each fading region for the discrete-power discrete-rate policy was chosen to have an average BER equal to the instantaneous BER of the discrete-rate continuous-power adaptive policy. We see from these figures that, for variable-rate MQAM with a small set of constellations, restricting the power to a single value for each constellation degrades spectral efficiency by about $1 - 2$ dB relative to continuous-power adaptation. For comparison, we also plot the maximum efficiency (9.17) for continuous-power and continuous-rate adaptation. All discrete-rate policies have performance that is within 3 dB of this theoretical maximum.

These figures also show the spectral efficiency of fixed-rate transmission with truncated channel inversion (9.26). The efficiency of this scheme is quite close to that of the discrete-power discrete-rate policy. However, to achieve this high efficiency, the optimal γ_0 is quite large, with a corresponding outage probability $P_{\rm out} = p(\gamma \leq \gamma_0)$ ranging from .1 to .6. Thus, this policy is similar to packet radio, with bursts of high-speed data when the channel conditions are favorable. The efficiency of total channel inversion (9.25) is also shown for log-normal shadowing; this efficiency equals zero in Rayleigh fading. We also plot the spectral efficiency of nonadaptive transmission, where both the transmission rate and power are constant. As discussed in Section 9.3.1, the average BER in this case is obtained by averaging the probability of error (9.31) over the fade distribution $p(\gamma)$. The spectral efficiency is obtained by determining the value of M that yields a 10^{-3} average BER for the given value of $\bar{\gamma}$, as illustrated in Example 9.3. Nonadaptive transmission clearly suffers a large spectral efficiency loss in exchange for its simplicity. However, if the channel varies rapidly and cannot be accurately estimated, nonadaptive

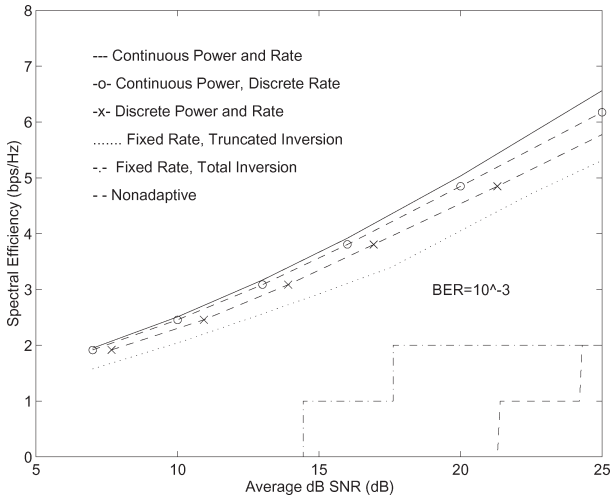


Figure 9.8: Efficiency in log-normal shadowing ($\sigma_{\psi_{\mathrm{dB}}} = 8$ dB).

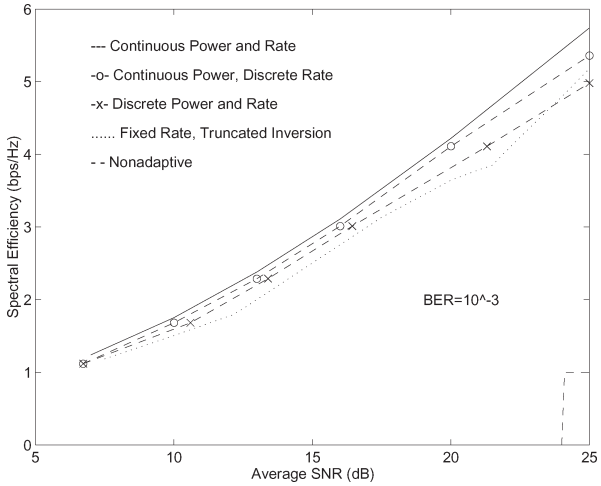


Figure 9.9: Efficiency in Rayleigh fading.

transmission may be the best alternative. Similar curves can be obtained for a target BER of 10^{-6}, with roughly the same spectral efficiency loss relative to a 10^{-3} BER as was exhibited in Figures 9.2 and 9.3.

9.3.5 Average Fade Region Duration

The choice of the number of regions to use in the adaptive policy will depend on how fast the channel is changing as well as on the hardware constraints, which dictate how many constellations are available to the transmitter and at what rate the transmitter can change its constellation and power. Channel estimation and feedback considerations along with hardware constraints may dictate that the constellation remains constant over tens or even hundreds of symbols. In addition, power-amplifier linearity requirements and out-of-band emission constraints may restrict the rate at which power can be adapted. An in-depth discussion of hardware implementation issues can be found in [18]. However, determining how long the SNR γ remains within a particular fading region R_j is of interest, since it determines the trade-off between the number of regions and the rate of power and constellation adaptation. We now investigate the time duration over which the SNR remains within a given fading region.

Table 9.2: Average fade region duration $\bar{\tau}_j$ for $f_D = 100$ Hz

Region(j)	$\bar{\gamma} = \mathbf{10}\ \boldsymbol{dB}$	$\bar{\gamma} = \mathbf{20}\ \boldsymbol{dB}$
0	2.23 ms	0.737 ms
1	0.83 ms	0.301 ms
2	3.00 ms	1.06 ms
3	2.83 ms	2.28 ms
4	1.43 ms	3.84 ms

Let $\bar{\tau}_j$ denote the average time duration that γ stays within the jth fading region. Let $A_j = \gamma_K^* M_j$ for γ_K^* and M_j as previously defined. The jth fading region is then defined as $\{\gamma : \ A_j \leq \gamma < A_{j+1}\}$. We call $\bar{\tau}_j$ the jth average fade region duration (AFRD). This definition is similar to the average fade duration (AFD; see Section 3.2.3), except that the AFD measures the average time that γ stays below a single level, whereas we are interested in the average time that γ stays between two levels. For the worst-case region ($j = 0$), these two definitions coincide.

Determining the exact value of $\bar{\tau}_j$ requires a complex derivation based on the joint density $p(\gamma, \dot{\gamma})$, and it remains an open problem. However, a good approximation can be obtained using the finite-state Markov model described in Section 3.2.4. In this model, fading is approximated as a discrete-time Markov process with time discretized to a given interval T, typically the symbol time. It is assumed (i) that the fade value γ remains within one region over a symbol period and (ii) that from a given region the process can only transition to the same region or to adjacent regions. Note that this approximation can lead to longer deep fade durations than more accurate models [19]. The transition probabilities between regions under this assumption are given as

$$p_{j,j+1} = \frac{L_{j+1}T}{\pi_j}, \quad p_{j,j-1} = \frac{L_j T}{\pi_j}, \quad p_{j,j} = 1 - p_{j,j+1} - p_{j,j-1}, \tag{9.27}$$

where L_j is the level-crossing rate at A_j and π_j is the steady-state distribution corresponding to the jth region: $\pi_j = p(A_j \leq \gamma < A_{j+1})$. Since the time over which the Markov process stays in a given state is geometrically distributed [20, Chap. 2.3], $\bar{\tau}_j$ is given by

$$\bar{\tau}_j = \frac{T}{p_{j,j+1} + p_{j,j-1}} = \frac{\pi_j}{L_{j+1} + L_j}. \tag{9.28}$$

The value of $\bar{\tau}_j$ is thus a simple function of the level crossing rate and the fading distribution. Whereas the level crossing rate is known for Rayleigh fading [21, Chap. 1.3.4], it cannot be obtained for log-normal shadowing because the joint distribution $p(\gamma, \dot{\gamma})$ for this fading type is unknown.

In Rayleigh fading, the level crossing rate is given by (3.55) as

$$L_j = \sqrt{\frac{2\pi A_j}{\bar{\gamma}}} f_D e^{-A_j/\bar{\gamma}}, \tag{9.29}$$

where $f_D = v/\lambda$ is the Doppler frequency. Substituting (9.29) into (9.28), we easily see that $\bar{\tau}_j$ is inversely proportional to the Doppler frequency. Moreover, π_j and A_j do not depend on f_D and so, if we compute $\bar{\tau}_j$ for a given Doppler frequency f_D, then we can compute $\bar{\hat{\tau}}_j$ corresponding to another Doppler frequency $\hat{f}_D$ as

$$\bar{\hat{\tau}}_j = \frac{f_D}{\hat{f}_D} \bar{\tau}_j. \tag{9.30}$$

Table 9.2 shows the $\bar{\tau}_j$ values corresponding to five regions ($M_j = 0, 2, 4, 16, 64$) in Rayleigh fading[2] for $f_D = 100$ Hz and two average power levels: $\bar{\gamma} = 10$ dB($\gamma_K^* = 1.22$) and $\bar{\gamma} = 20$ dB($\gamma_K^* = 1.685$). The AFRD for other Doppler frequencies is easily obtained using the table values and (9.30). This table indicates that, even at high velocities, for rates of 100 kilosymbols/second the discrete-rate discrete-power policy will maintain the same constellation and transmit power over tens to hundreds of symbols.

Example 9.4: Find the AFRDs for a Rayleigh fading channel with $\bar{\gamma} = 10$ dB, $M_j = 0, 2, 4, 16, 64, 64$, and $f_D = 50$ Hz.
Solution: We first note that all parameters are the same as used in the calculation of Table 9.2 except that the Doppler $\hat{f}_D = 50$ Hz is half the Doppler of $f_D = 100$ Hz used to compute the table values. Thus, from (9.30), we obtain the AFRDs with this new Doppler by multiplying each value in the table by $f_D/\hat{f}_D = 2$.

In shadow fading we can obtain a coarse approximation of $\bar{\tau}_j$ based on the shadowing autocorrelation function (2.44): $A(\delta) = \sigma^2_{\psi_{\rm dB}} e^{-\delta/X_c}$, where $\delta = v\tau$ for v the mobile's velocity. Specifically, we can approximate the AFRD for all regions as $\bar{\tau}_j \approx .1X_c/v$, since then the correlation between fade levels separated in time by $\bar{\tau}_j$ is .9. Thus, for a small number of regions it is likely that γ will remain within the same region over this time period.

9.3.6 Exact versus Approximate Bit Error Probability

The adaptive policies described in prior sections are based on the BER upper bounds of Section 9.3.1. Since these are upper bounds, they will lead to a lower BER than the target. We would like to see how the BER achieved with these policies differs from the target BER. A more accurate value for the BER achieved with these policies can be obtained by simulation or by using a better approximation for BER than the upper bounds. From Table 6.1, the BER of MQAM with Gray coding at high SNRs is well approximated by

$$P_b \approx \frac{4}{\log_2 M} Q\left(\sqrt{\frac{3\gamma}{M-1}}\right). \tag{9.31}$$

Moreover, for the continuous-power discrete-rate policy, $\gamma = E_s/N_0$ for the jth signal constellation is

$$\frac{E_s(j)}{N_0} = \frac{M_j - 1}{K}. \tag{9.32}$$

Thus, we can obtain a more accurate analytical expression for the average BER associated with our adaptive policies by averaging over the BER (9.31) for each signal constellation as

$$\bar{P}_b = \frac{\sum_{j=1}^{N-1} 4Q(\sqrt{3/K}) \int_{\gamma_K^* M_j}^{\gamma_K^* M_{j+1}} p(\gamma)\, d\gamma}{\sum_{j=1}^{N-1} \log_2 M_j \int_{\gamma_K^* M_j}^{\gamma_K^* M_{j+1}} p(\gamma)\, d\gamma} \tag{9.33}$$

with $M_N = \infty$.

We plot the analytical expression (9.33) along with the simulated BER for the variable rate and power MQAM with a target BER of 10^{-3} in Figures 9.10 and 9.11 for log-normal shadowing and Rayleigh fading, respectively. The simulated BER is slightly better than the analytical calculation of (9.33) because this equation is based on the nearest neighbor bound for the maximum number of nearest neighbors. Both the simulated and analytical BER are smaller than the target BER of 10^{-3} for $\bar{\gamma} > 10$ dB. The BER bound of 10^{-3} breaks down at low SNRs, since (9.7) is not applicable to BPSK, so we must use the looser bound (9.6). Since the adaptive policy often uses the

[2]The validity of the finite-state Markov model for Rayleigh fading channels has been investigated in [22].

BPSK constellation at low SNRs, the P_b will be larger than that predicted from the tight bound (9.7). That the simulated BER is less than the target at high SNRs implies that the analytical calculations in Figures 9.6 and 9.7 are pessimistic; a slightly higher efficiency could be achieved while still maintaining the target P_b of 10^{-3}.

9.3.7 Channel Estimation Error and Delay

In this section we examine the effects of estimation error and delay, where the estimation error $\varepsilon = \hat{\gamma}/\gamma \neq 1$ and the delay $i_d = i_f + i_e \neq 0$. We first consider the estimation error. Suppose the transmitter adapts its power and rate relative to a target BER P_b^t based on the channel estimate $\hat{\gamma}$ instead of the true value γ. By (9.8), the BER is then bounded by

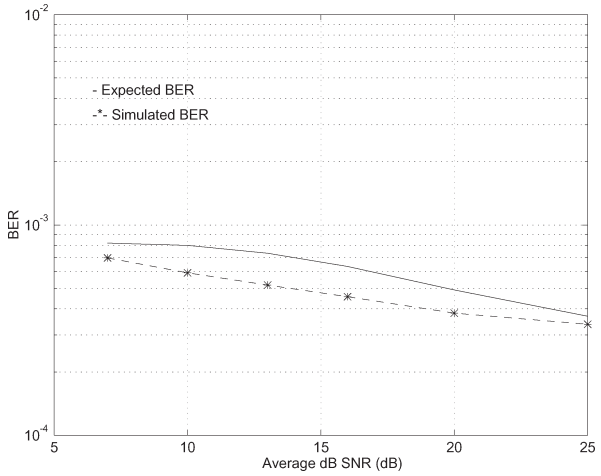


Figure 9.10: BER for log-normal shadowing (six regions).

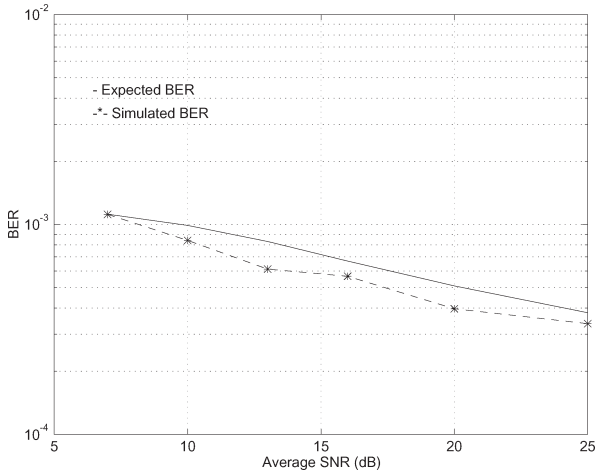


Figure 9.11: BER for Rayleigh fading (five regions).

$$P_b(\gamma, \hat{\gamma}) \leq .2 \exp\left[\frac{-1.5\gamma}{M(\hat{\gamma}) - 1} \frac{P(\hat{\gamma})}{\bar{P}}\right] = .2[5P_b^t]^{1/\varepsilon}, \tag{9.34}$$

where the right-hand equality is obtained by substituting the optimal rate (9.9) and power (9.13) policies. For $\varepsilon = 1$, (9.34) reduces to the target BER P_b^t. For $\varepsilon \neq 1$, $\varepsilon > 1$ yields an increase in BER above the target, and $\varepsilon < 1$ yields a decrease in BER.

The average probability of bit error is given by

$$\bar{P}_b = \frac{\mathbf{E}[\text{number of bits in error per transmission}]}{\mathbf{E}[\text{number of bits per transmission}]}. \tag{9.35}$$

Given the BER bound (9.34) for constellation $M(\hat{\gamma})$ with rate $\log_2 M(\hat{\gamma})$, we can bound the average probability of error (9.35) of adaptive modulation given channel estimation error and delay as

$$\bar{P}_b = \frac{\int_0^\infty \int_{\gamma_K}^\infty P_b(\gamma,\hat{\gamma}) M(\hat{\gamma}) p(\gamma,\hat{\gamma})\, d\hat{\gamma} d\gamma}{\int_0^\infty M(\hat{\gamma}) p(\hat{\gamma})\, d\hat{\gamma}} \leq \frac{\int_0^\infty \int_{\gamma_K}^\infty .2[.5P_b^t]^{\gamma/\hat{\gamma}} M(\hat{\gamma}) p(\gamma,\hat{\gamma})\, d\hat{\gamma} d\gamma}{\int_0^\infty M(\hat{\gamma}) p(\hat{\gamma})\, d\hat{\gamma}}. \tag{9.36}$$

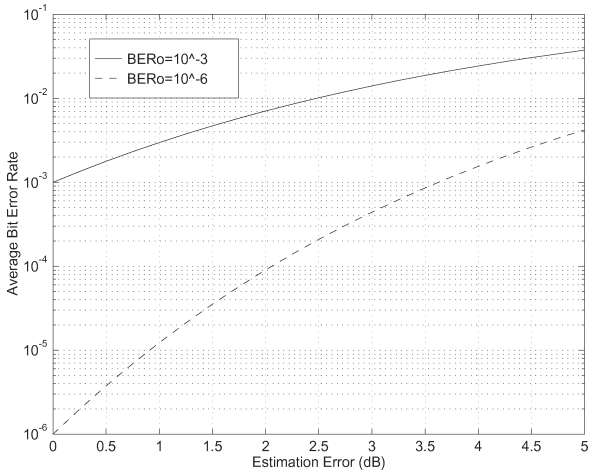


Figure 9.12: Effect of estimation error on BER.

The joint distribution $p(\gamma,\hat{\gamma})$ depends on the channel estimation technique. It has been shown [23] that, when the channel is estimated using pilot symbols, the joint distribution of the signal envelope and its estimate is bivariate Rayleigh. This joint distribution was used in [23] to obtain the probability of error for nonadaptive modulation with channel estimation errors. This analysis can be extended to adaptive modulation using a similar methodology.

If the estimation error stays within some finite range then we can bound its effect using (9.34). We plot the BER increase as a function of a constant ε in Figure 9.12. This figure shows that for a target BER of 10^{-3} the estimation error should be less than 1 dB, and for a target BER of 10^{-6} it should be less than .5 dB. These values are pessimistic, since they assume a constant value of estimation error. Even so, the estimation error can be kept within this range using the pilot-symbol assisted estimation technique described in [24] with appropriate choice of parameters. When the channel is underestimated $(\varepsilon < 1)$, the BER decreases but there will also be some loss in spectral efficiency, since the mean of the channel estimate $\bar{\hat{\gamma}}$ will differ from the true mean $\bar{\gamma}$. The effect of this average power estimation error is characterized in [25].

Suppose now that the channel is estimated perfectly $(\varepsilon = 1)$ but the delay i_d of the estimation and feedback path is nonzero. Thus, at time i the transmitter will use the delayed version of the channel estimate $\hat{\gamma}[i] = \gamma[i - i_d]$ to adjust its power and rate. It was shown in [26] that, conditioned on the outdated channel estimates, the received signal follows a Rician distribution, and the probability of error can then be computed by averaging over the distribution of the estimates. Moreover, [26] develops adaptive coding designs to mitigate the effect of estimation

delay on the performance of adaptive modulation. Alternatively, channel prediction can be used to mitigate these effects [27].

The increase in BER from estimation delay can also be examined in the same manner as in (9.34). Given a target BER P_b^t, the exact channel SNR $\gamma[i]$, and its delayed value $\gamma[i - i_d]$, we have

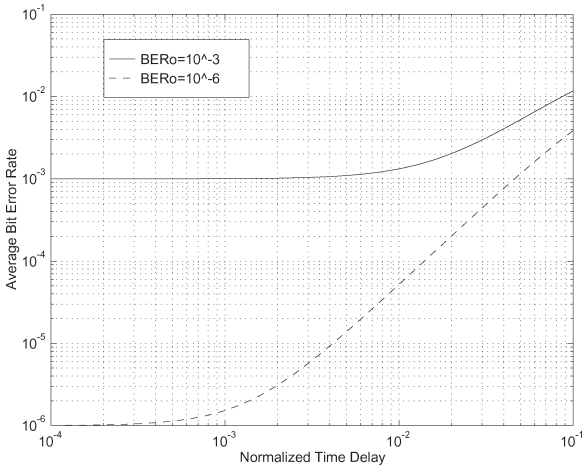


Figure 9.13: Effect of normalized delay $(i_d f_D)$ on BER.

$$P_b(\gamma[i],\ \gamma[i-i_d]) \leq .2 \exp\left[\frac{-1.5\gamma[i]}{M(\gamma[i-i_d])-1}\frac{P(\gamma[i-i_d])}{\bar{P}}\right]$$
$$= .2[5P_b^t]^{\gamma[i]/\gamma[i-i_d]}. \tag{9.37}$$

Define $\xi[i,\ i_d] = \gamma[i]/\gamma[i - i_d]$. Since $\gamma[i]$ is stationary and ergodic, the distribution of $\xi[i,\ i_d]$ conditioned on $\gamma[i]$ depends only on i_d and the value of $\gamma = \gamma[i]$. We denote this distribution by $p_{i_d}(\xi|\gamma)$. The instantaneous data rate $\log_2 M(\gamma[i - i_d])$ is based on the value of $\gamma[i - i_d]$ but is transmitted over the channel at time i. So the time delay i_d does not impact the average data rate. However, this delay impacts the instantaneous BER since there is a mismatch between the constellation to be transmitted at time i, $M(\gamma[i])$, and associated power control to achieve the target BER P_b^t, and the constellation that is transmitted at that time, $M(\gamma[i - i_d])$. As shown in [28], the average BER is obtained by integrating over the distribution of this instantaneous BER which depends only on ξ and γ:

$$P_b[i_d] = \int_{\gamma_K}^{\infty}\left[\int_0^{\infty} .2[5P_b^t]^{\xi} p_{i_d}(\xi|\gamma)d\xi\right] p(\gamma)\, d\gamma, \tag{9.38}$$

where γ_K is the cutoff level of the optimal policy and $p(\gamma)$ is the fading distribution. The outer integral is taken from γ_K rather than zero since we only transmit over the channel when $\gamma \geq \gamma_K$. The distribution $p_{i_d}(\xi|\gamma)$ will depend on the autocorrelation of the fading process. A closed-form expression for $p_{i_d}(\xi|\gamma)$ in Nakagami fading (of which Rayleigh fading is a special case) is derived in [28]. Using this distribution in (9.38), we obtain the average BER in Rayleigh fading as a function of the delay parameter i_d. A plot of (9.38) versus the normalized time delay $i_d f_D$ is shown in Figure 9.13. From this figure we see that the total estimation and feedback path delay must be kept to within $.001/f_D$ in order to keep the BER near its desired target.

9.3.8 Adaptive Coded Modulation

Additional coding gain can be achieved with adaptive modulation by superimposing trellis or lattice codes on the adaptive modulation. Specifically, by using the subset partitioning inherent to coded modulation, trellis (or lattice) codes designed for AWGN channels can be superimposed directly onto the adaptive modulation with the same approximate coding gain. The basic idea of adaptive coded modulation is to exploit the separability of code and constellation design that is characteristic of coded modulation, as described in Section 8.9.

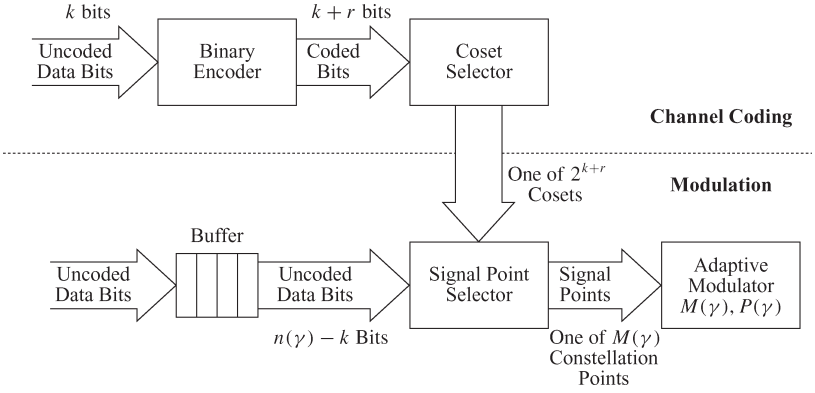


Figure 9.14: Adaptive coded modulation scheme.

Coded modulation is a natural coding scheme to use with variable-rate variable-power MQAM, since the channel coding gain is essentially independent of the modulation. We can therefore adjust the power and rate (number of levels or signal points) in the transmit constellation relative to the instantaneous SNR *without* affecting the channel coding gain, as we now describe in more detail.

The coded modulation scheme is shown in Figure 9.14. The coset code design is the same as it would be for an AWGN channel; that is, the lattice structure and conventional encoder follow the trellis or lattice coded modulation designs outlined in Section 8.9. Let G_c denote the coding gain of the coded modulation, as given by (8.95). The modulation works as follows. The signal constellation is a square lattice with an adjustable number of constellation points M. The size of the MQAM signal constellation from which the signal point is selected is determined by the transmit power, which is adjusted relative to the instantaneous SNR and the desired BER, as in the uncoded case described in Section 9.3.2.

Specifically, if the BER approximation (9.7) is adjusted for the coding gain, then for a particular SNR $= \gamma$ we have

$$P_b \approx .2e^{-1.5(\gamma G_c/M-1)}, \tag{9.39}$$

where M is the size of the transmit signal constellation. As in the uncoded case, using the tight bound (9.7) allows us to adjust the number of constellation points M and signal power relative to the instantaneous SNR in order to maintain a fixed BER:

$$M(\gamma) = 1 + \frac{1.5\gamma G_c}{-\ln(5P_b)}\frac{P(\gamma)}{\bar{P}}. \tag{9.40}$$

The number of uncoded bits required to select the coset point is

$$n(\gamma) - 2k/N = \log_2 M(\gamma) - 2(k+r)/N.$$

Since this value varies with time, these uncoded bits must be queued until needed, as shown in Figure 9.14.

The bit rate per transmission is $\log_2 M(\gamma)$, and the data rate is $\log_2 M(\gamma) - 2r/N$. Therefore, we maximize the data rate by maximizing $\mathbf{E}[\log_2 M]$ relative to the average power constraint. From this maximization, we obtain the optimal power adaptation policy for this modulation scheme:

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} 1/\gamma_0 - 1/\gamma K_c & \gamma \geq \gamma_0/K_c, \\ 0 & \gamma < \gamma_0/K_c, \end{cases} \tag{9.41}$$

where $\gamma_{K_c} = \gamma_0/K_c$ is the cutoff fade depth for $K_c = KG_c$ with K given by (9.10). This is the same as the optimal policy for the uncoded case (9.13) with K replaced by K_c. Thus, the coded modulation increases the effective transmit power by G_c relative to the uncoded variable-rate variable-power MQAM performance. The adaptive data rate is obtained by substituting (9.41) into (9.40) to get

$$M(\gamma) = \left(\frac{\gamma}{\gamma_{K_c}}\right). \tag{9.42}$$

The resulting spectral efficiency is

$$\frac{R}{B} = \int_{\gamma K_C}^{\infty} \log_2 \left(\frac{\gamma}{\gamma_{K_c}}\right) p(\gamma)\, d\gamma. \tag{9.43}$$

If the constellation expansion factor is not included in the coding gain G_c, then we must subtract $2r/N$ from (9.43) to get the data rate. More details on this adaptive coded modulation scheme can be found in [29], along with plots of the spectral efficiency for adaptive trellis coded modulation of varying complexity. These results indicate that adaptive trellis coded modulation can achieve within 5 dB of Shannon capacity at reasonable complexity and that the coding gains of superimposing a given trellis code onto uncoded adaptive modulation are roughly equal to the coding gains of the trellis code in an AWGN channel.

9.4 General M-ary Modulations

The variable rate and power techniques already described for MQAM can be applied to other M-ary modulations. For any modulation, the basic premise is the same: the transmit power and constellation size are adapted to maintain a given fixed instantaneous BER for each symbol while maximizing average data rate. In this section we will consider optimal rate and power adaptation for both continuous-rate and discrete-rate variation of general M-ary modulations.

9.4.1 Continuous-Rate Adaptation

We first consider the case where both rate and power can be adapted continuously. We want to find the optimal power $P(\gamma)$ and rate $k(\gamma) = \log_2 M(\gamma)$ adaptation for general M-ary modulation that maximizes the average data rate $\mathbf{E}[k(\gamma)]$ with average power $\bar{P}$ while meeting a given BER target. This optimization is simplified when the exact or approximate probability of bit error for the modulation can be written in the following form:

$$P_b(\gamma) \approx c_1 \exp\left[\frac{-c_2\gamma(P(\gamma)/\bar{P})}{2^{c_3 k(\gamma)} - c_4}\right], \tag{9.44}$$

where c_1, c_2, and c_3 are positive fixed constants and c_4 is a real constant. For example, in the BER bounds for MQAM given by (9.6) and (9.7), $c_1 = 2$ or .2, $c_2 = 1.5, c_3 = 1$, and $c_4 = 1$. The probability of bit error for most M-ary modulations can be approximated in this form with appropriate curve fitting.

The advantage of (9.44) is that, when $P_b(\gamma)$ is in this form, we can invert it to express the rate $k(\gamma)$ as a function of the power adaptation $P(\gamma)$ and the BER target P_b as follows:

$$k(\gamma) = \log_2 M(\gamma) = \begin{cases} \frac{1}{c_3}\log_2\left[c_4 - \frac{c_2\gamma}{\ln(P_b/c_1)}\frac{P(\gamma)}{\bar{P}}\right] & P(\gamma) \geq 0,\ k(\gamma) \geq 0, \\ 0 & \text{else.} \end{cases} \tag{9.45}$$

To find the power and rate adaptation that maximize spectral efficiency $\mathbf{E}[k(\gamma)]$, we create the Lagrangian

$$J(P(\gamma)) = \int_0^\infty k(\gamma)p(\gamma)d\gamma + \lambda\left[\int_0^\infty P(\gamma)p(\gamma)d\gamma - \bar{P}\right]. \tag{9.46}$$

The optimal adaptation policy maximizes this Lagrangian with nonnegative rate and power, so it satisfies

$$\frac{\partial J}{\partial P(\gamma)} = 0,\ P(\gamma) \geq 0,\ k(\gamma) \geq 0. \tag{9.47}$$

Solving (9.47) for $P(\gamma)$ with (9.45) for $k(\gamma)$ yields the optimal power adaptation

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} -1/(c_3(\ln 2)\lambda\bar{P}) - 1/(\gamma K) & P(\gamma) \geq 0,\ k(\gamma) \geq 0, \\ 0 & \text{else,} \end{cases} \tag{9.48}$$

where

$$K = -\frac{c_2}{c_4\ln(P_b/c_1)}. \tag{9.49}$$

The power adaptation (9.48) can be written in the more simplified form

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \mu - 1/(\gamma K) & P(\gamma) \geq 0,\ k(\gamma) \geq 0, \\ 0 & \text{else.} \end{cases} \tag{9.50}$$

The constant μ in (9.50) is determined from the average power constraint (9.12).

Although the analytical expression for the optimal power adaptation (9.50) looks simple, its behavior is highly dependent on the c_4 values in the P_b approximation (9.44). For (9.44) given by the MQAM approximations (9.6) or (9.7), the power adaptation is the water-filling formula given by (9.13). However, water-filling is not optimal in all cases, as we now show.

Based on (6.21), with Gray coding the BER for MPSK is tightly approximated as

$$P_b \approx \frac{2}{\log_2 M}Q\left(\sqrt{2\gamma}\sin\left(\frac{\pi}{M}\right)\right). \tag{9.51}$$

However, (9.51) is not in the desired form (9.44). In particular, the Q-function is not easily inverted to obtain the optimal rate and power adaptation for a given target BER. Let us therefore consider the following three P_b bounds for MPSK, which are valid for $k(\gamma) \geq 2$.

$$\text{Bound 1:}\ P_b(\gamma) \approx .05\exp\left[\frac{-6\gamma(P(\gamma)/\bar{P})}{2^{1.9k(\gamma)} - 1}\right]. \tag{9.52}$$

$$\text{Bound 2:}\ P_b(\gamma) \approx .2\exp\left[\frac{-7\gamma(P(\gamma)/\bar{P})}{2^{1.9k(\gamma)} + 1}\right]. \tag{9.53}$$

$$\text{Bound 3:}\ P_b(\gamma) \approx .25\exp\left[\frac{-8\gamma(P(\gamma)/\bar{P})}{2^{1.94k(\gamma)}}\right]. \tag{9.54}$$

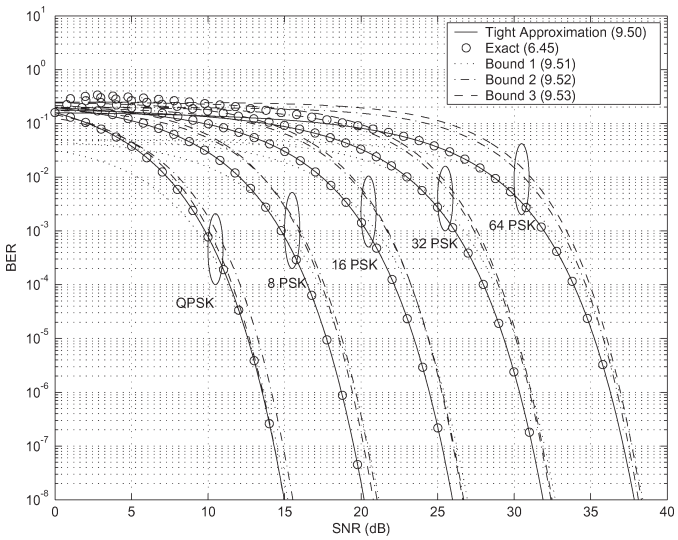


Figure 9.15: BER bounds for MPSK.

The bounds are plotted in Figure 9.15 along with the tight approximation (9.51). We see that all bounds well approximate the exact BER given by (6.55), especially at high SNRs.

In the first bound (9.52), $c_1 = .05, c_2 = 6, c_3 = 1.9$, and $c_4 = 1$. Thus, in (9.50), $K = -c_2/(c_4 \ln(P_b/c_1))$ is positive as long as the target P_b is less than .05, which we assume. Therefore μ must be positive for the power adaptation $P(\gamma)/\bar{P} = \mu - 1/(\gamma K)$ to be positive about a cutoff SNR γ_0. Moreover, for K positive, $k(\gamma) \geq 0$ for any $P(\gamma) \geq 0$. Thus, with μ and $k(\gamma)$ positive, (9.50) can be expressed as

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} 1/(\gamma_0 K) - 1/(\gamma K) & P(\gamma) \geq 0, \\ 0 & \text{else}, \end{cases} \tag{9.55}$$

where $\gamma_0 \geq 0$ is a cutoff fade depth below which no signal is transmitted. Like μ, this cutoff value is determined by the average power constraint (9.12). The power adaptation (9.55) is the same water-filling as in adaptive MQAM given by (9.13), which results from the similarity of the MQAM P_b bounds (9.7) and (9.6) to the MPSK bound (9.52). The corresponding optimal rate adaptation, obtained by substituting (9.55) into (9.45), is

$$k(\gamma) = \begin{cases} (1/c_3)\log_2(\gamma/\gamma_0) & \gamma \geq \gamma_0, \\ 0 & \text{else}, \end{cases} \tag{9.56}$$

which is also in the same form as the adaptive MQAM rate adaptation (9.16).

Let us now consider the second bound (9.53). Here $c_1 = .2, c_2 = 7, c_3 = 1.9$, and $c_4 = -1$. Thus, $K = -c_2/(c_4 \ln(P_b/c_1))$ is negative for a target $P_b < .2$, which we assume. From (9.45), with K negative we must have $\mu \geq 0$ in (9.50) to make $k(\gamma) \geq 0$. Then the optimal power adaptation such that $P(\gamma) \geq 0$ and $k(\gamma) \geq 0$ becomes

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \mu - 1/(\gamma K) & k(\gamma) \geq 0, \\ 0 & \text{else}. \end{cases} \tag{9.57}$$

From (9.45), the optimal rate adaptation then becomes

$$k(\gamma) = \begin{cases} (1/c_3)\log_2(\gamma/\gamma_0) & \gamma \geq \gamma_0, \\ 0 & \text{else}, \end{cases} \tag{9.58}$$

where $\gamma_0 = -1/(K\mu)$ is a cutoff fade depth below which the channel is not used. Note that for the first bound (9.52) the positivity constraint on power ($P(\gamma) \geq 0$) dictates the cutoff fade depth, whereas for this bound (9.53) the positivity constraint on rate ($k(\gamma) \geq 0$) determines the cutoff. We can rewrite (9.57) in terms of γ_0 as

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} -1/(\gamma_0 K) - 1/(\gamma K) & \gamma \geq \gamma_0, \\ 0 & \text{else}. \end{cases} \tag{9.59}$$

This power adaptation is an *inverse water-filling*: since K is negative, less power is used as the channel SNR increases above the optimized cutoff fade depth γ_0. As usual, the value of γ_0 is obtained based on the average power constraint (9.12).

Finally, for the third bound (9.54), $c_1 = .25, c_2 = 8, c_3 = 1.94$, and $c_4 = 0$. Thus, $K = -c_2/(c_4 \ln(P_b/c_1)) = \infty$ for a target $P_b < .25$, which we assume. From (9.50), the optimal power adaptation becomes

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \mu & k(\gamma) \geq 0,\ P(\gamma) \geq 0, \\ 0 & \text{else}. \end{cases} \tag{9.60}$$

This is *on-off* power transmission: power is either zero or a constant nonzero value. By (9.45), the optimal rate adaptation $k(\gamma)$ with this power adaptation is

$$k(\gamma) = \begin{cases} (1/c_3)\log_2(\gamma/\gamma_0) & \gamma \geq \gamma_0, \\ 0 & \text{else}, \end{cases} \tag{9.61}$$

where $\gamma_0 = -\ln(P_b/c_1)/(c_2\mu)$ is a cutoff fade depth below which the channel is not used. As for the previous bound, it is the rate positivity constraint that determines the cutoff fade depth γ_0. The optimal power adaptation as a function of γ_0 is

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} K_0/\gamma_0 & \gamma \geq \gamma_0, \\ 0 & \text{else}, \end{cases} \tag{9.62}$$

where $K_0 = -\ln(P_b/c_1)/c_2$. The value of γ_0 is determined from the average power constraint to satisfy

$$\frac{K_0}{\gamma_0}\int_{\gamma_0}^{\infty} p(\gamma)\, d\gamma = 1. \tag{9.63}$$

Thus, for all three P_b approximations in MPSK, the optimal adaptive rate schemes (9.56), (9.58), and (9.61) have the same form whereas the optimal adaptive power schemes (9.55), (9.59), and (9.62) have different forms. The optimal power adaptations (9.55), (9.59), and (9.62) are plotted in Figure 9.16 for Rayleigh fading with a target BER of 10^{-3} and $\bar{\gamma} = 30$ dB. This figure clearly shows the water-filling, inverse water-filling, and on–off behavior of the different schemes. Note that the cutoff γ_0 for all these schemes is roughly the same. We also see from this figure that even though the power adaptation schemes are different at low SNRs, they are almost the same at high SNRs. Specifically, we see that for $\gamma < 10\,\text{dB}$, the optimal transmit power adaptations are dramatically different, whereas for $\gamma \geq 10\,\text{dB}$ they rapidly converge to the same constant value. From the cumulative distribution function of γ, also shown in Figure 9.16, the probability that γ is less than 10 is 0.01. Thus, although the optimal power adaptation corresponding to low SNRs is very different for the different techniques, this behavior has little impact on spectral efficiency because the probability of being at those low SNRs is quite small.

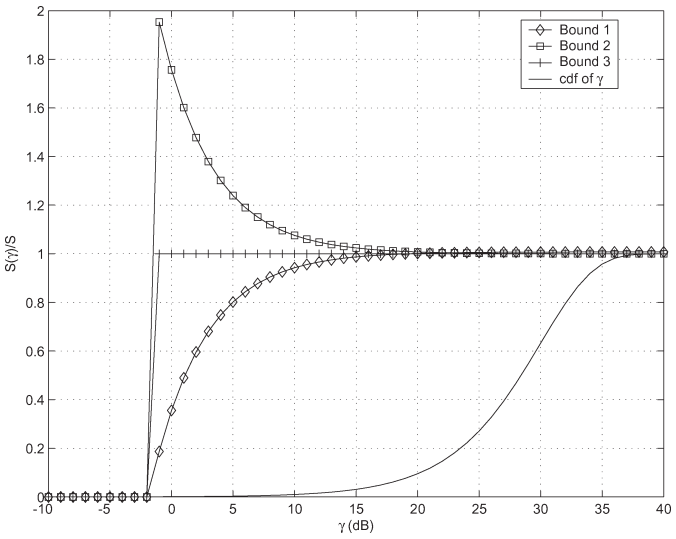


Figure 9.16: Power adaptation for MPSK BER bounds (Rayleigh fading, $P_{\rm b} = 10^{-3}, \bar{\gamma} = 30\,{\rm dB}$).

9.4.2 Discrete-Rate Adaptation

We now assume a given discrete set of constellations $\mathcal{M} = \{M_0 = 0,\ \ldots,\ M_{N-1}\}$, where M_0 corresponds to no data transmission. The rate corresponding to each of these constellations is $k_j = \log_2 M_j$, $j = 0, \ldots, N-1$, where $k_0 = 0$. Each rate k_j, $j > 0$, is assigned to a fading region of γ-values $R_j = [\gamma_{j-1}, \gamma_j)$, $j = 0, \ldots, N-1$, for $\gamma_{-1} = 0$ and $\gamma_{N-1} = \infty$. The boundaries γ_j, $j = 0, \ldots, N-2$, are optimized as part of the adaptive policy. The channel is not used for $\gamma < \gamma_0$. We again assume that P_b is approximated using the general formula (9.44). Then the power adaptation that maintains the target BER above the cutoff γ_0 is

$$\frac{P(\gamma)}{\bar{P}} = \frac{h(k_j)}{\gamma},\ \gamma_{j-1} \le \gamma \le \gamma_j, \tag{9.64}$$

where

$$h(k_j) = -\frac{\ln(P_b/c_1)}{c_2}(2^{c_3 k_j} - c_4). \tag{9.65}$$

The region boundaries $\gamma_0, \ldots, \gamma_{N-2}$ that maximize spectral efficiency are found using the Lagrange equation

$$J(\gamma_0,\ \gamma_1, \ldots, \gamma_{N-2}) = \sum_{j=1}^{N-1} k_j \int_{\gamma_{j-1}}^{\gamma_j} p(\gamma)\, d\gamma + \lambda \left[\sum_{j=1}^{N-1} \int_{\gamma_{j-1}}^{\gamma_j} \frac{h(k_j)}{\gamma} p(\gamma) d\gamma - 1 \right]. \tag{9.66}$$

The optimal rate region boundaries are obtained by solving the following equation for γ_j:

$$\frac{\partial J}{\partial \gamma_j} = 0,\ 0 \le j \le N-2. \tag{9.67}$$

This yields

$$\gamma_0 = \frac{h(k_1)}{k_1}\rho \tag{9.68}$$

and

$$\gamma_j = \frac{h(k_{j+1}) - h(k_j)}{k_{j+1} - k_j}\rho,\ 1 \le i \le N-2, \tag{9.69}$$

where ρ is determined by the average power constraint

$$\sum_{j=1}^{N-1} \int_{\gamma_{j-1}}^{\gamma j} \frac{h(k_j)}{\gamma} p(\gamma)\, d\gamma = 1. \tag{9.70}$$

9.4.3 Average BER Target

Suppose now that we relax our assumption that the P_b target must be met on every symbol transmission, requiring instead that just the average P_b be below some target average $\bar{P}_b$. In this case, in addition to adapting rate and power, we can also adapt the instantaneous $P_b(\gamma)$ subject to the average constraint $\bar{P}_b$. This gives an additional degree of freedom in adaptation that may lead to higher spectral efficiencies. The average probability of error for adaptive modulation is given in (9.35).
When the bit rate $k(\gamma)$ is continuously adapted this becomes

$$\bar{P}_b = \frac{\int_0^\infty P_b(\gamma)k(\gamma)p(\gamma)\, d\gamma}{\int_0^\infty k(\gamma)p(\gamma)\, d\gamma}, \tag{9.71}$$

and when $k(\gamma)$ takes values in a discrete set this becomes

$$\bar{P}_b = \frac{\sum_{j=1}^{N-1} k_j \int_{\gamma_{j-1}}^{\gamma_j} P_b(\gamma)p(\gamma)\, d\gamma}{\sum_{j=1}^{N-1} k_j \int_{\gamma_{j-1}}^{\gamma_j} p(\gamma)\, d\gamma}. \tag{9.72}$$

We now derive the optimal continuous rate, power, and BER adaptation to maximize spectral efficiency $\mathbf{E}[k(\gamma)]$ subject to an average power constraint $\bar{P}$ and the average BER constraint (9.71). As with the instantaneous BER constraint, this is a standard constrained optimization problem, which we solve using the Lagrange method. We now require two Lagrangians for the two constraints: average power and average BER. Specifically, the Lagrange equation is

$$\begin{aligned} J(k(\gamma),\ P(\gamma)) = &\int_0^\infty k(\gamma)p(\gamma)\, d\gamma \\ &+ \lambda_1 \left[\int_0^\infty P_b(\gamma)k(\gamma)p(\gamma)\, d\gamma - \bar{P}_b \int_0^\infty k(\gamma)p(\gamma)\, d\gamma \right] \\ &+ \lambda_2 \left[\int_0^\infty P(\gamma)p(\gamma)\, d\gamma - \bar{P} \right]. \end{aligned} \tag{9.73}$$

The optimal rate and power adaptation must satisfy

$$\frac{\partial J}{\partial k(\gamma)} = 0 \text{ and } \frac{\partial J}{\partial P(\gamma)} = 0, \tag{9.74}$$

with the additional constraint that $k(\gamma)$ and $P(\gamma)$ be nonnegative for all γ.

Assume that P_b is approximated using the general formula (9.44). Define

$$f(k(\gamma)) = 2^{c_3 k(\gamma)} - c_4. \tag{9.75}$$

Then, using (9.44) in (9.73) and solving (9.74), we obtain that the power and BER adaptation that maximize spectral efficiency satisfy

$$\frac{P(\gamma)}{\bar{P}} = \max\left[\frac{f(k(\gamma))}{\frac{\partial f(k(\gamma))}{\partial k(\gamma)}}\lambda_2\bar{P}(\lambda_1\bar{P}_b - 1) - \frac{f(k(\gamma))^2}{c_2\gamma\frac{\partial f(k(\gamma))}{\partial k(\gamma)}k(\gamma)}, 0\right] \tag{9.76}$$

for nonnegative $k(\gamma)$ and

$$P_b(\gamma) = \frac{\lambda_2\bar{P}f(k(\gamma))}{\lambda_1 c_2\gamma k(\gamma)}. \tag{9.77}$$

Moreover, from (9.44), (9.76), and (9.77) we get that the optimal rate adaptation $k(\gamma)$ is either zero or the nonnegative solution of

$$\frac{\lambda_1\bar{P}_b - 1}{\frac{\partial f(k(\gamma))}{\partial k(\gamma)}\lambda_2\bar{P}} - \frac{f(k(\gamma))}{c_2\gamma\frac{\partial f(k(\gamma))}{\partial k(\gamma)}k(\gamma)} = \frac{1}{\gamma c_2}\ln\left[\frac{\lambda_1 c_1 c_2\gamma k(\gamma)}{\lambda_2\bar{P}f(k(\gamma))}\right]. \tag{9.78}$$

The values of $k(\gamma)$ and the Lagrangians λ_1 and λ_2 must be found through a numerical search whereby the average power constraint $\bar{P}$ and average BER constraint (9.71) are satisfied.

In the discrete-rate case, the rate is varied within a fixed set $k_0, \ldots, k_{N-1}$, where k_0 corresponds to no data transmission. We must determine region boundaries $\gamma_0, \ldots, \gamma_{N-2}$ such that we assign rate k_j to the rate region $[\gamma_{j-1},\ \gamma_j)$, where $\gamma_{-1} = 0$ and $\gamma_{N-1} = \infty$. Under this rate assignment we wish to maximize spectral efficiency through optimal rate, power, and BER adaptation subject to an average power and BER constraint. Since the set of possible rates and their corresponding rate region assignments are fixed, the optimal rate adaptation corresponds to finding the optimal rate region boundaries γ_j, $j = 0, \ldots, N-2$. The Lagrangian for this constrained optimization problem is

$$\begin{aligned} J(\gamma_0, \gamma_1, \ldots, \gamma_{N-2},\ P(\gamma)) = &\sum_{j=1}^{N-1} k_j \int_{\gamma_{j-1}}^{\gamma_j} p(\gamma)\, d\gamma \\ &+ \lambda_1\left[\sum_{j=1}^{N-1} k_j \int_{\gamma_{j-1}}^{\gamma_j} (P_b(\gamma) - \bar{P}_b)p(\gamma)\, d\gamma\right] \\ &+ \lambda_2\left[\int_{\gamma_0}^{\infty} P(\gamma)p(\gamma)\, d\gamma - \bar{P}\right]. \end{aligned} \tag{9.79}$$

The optimal power adaptation is obtained by solving the following equation for $P(\gamma)$:

$$\frac{\partial J}{\partial P(\gamma)} = 0. \tag{9.80}$$

Similarly, the optimal rate region boundaries are obtained by solving the following set of equations for γ_j:

$$\frac{\partial J}{\partial \gamma_j} = 0,\ 0 \le j \le N-2. \tag{9.81}$$

From (9.80) we see that the optimal power and BER adaptation must satisfy

$$\frac{\partial P_b(\gamma)}{\partial P(\gamma)} = \frac{-\lambda_2}{k_j\lambda_1},\ \gamma_{j-1} \le \gamma \le \gamma_j. \tag{9.82}$$

Substituting (9.44) into (9.82), we get that

$$P_b(\gamma) = \lambda \frac{f(k_j)}{\gamma k_j},\ \gamma_{j-1} \le \gamma \le \gamma_j, \tag{9.83}$$

where $\lambda = \bar{P}\lambda_2/(c_2\lambda_1)$. This form of BER adaptation is similar to the water-filling power adaptation: the instantaneous BER decreases as the channel quality improves. Now, setting the BER in (9.44) equal to (9.83) and solving for $P(\gamma)$ yields

$$P(\gamma) = P_j(\gamma),\ \gamma_{j-1} \le \gamma \le \gamma_j, \tag{9.84}$$

where

$$\frac{P_j(\gamma)}{\bar{P}} = \ln\left[\frac{\lambda f(k_j)}{c_1 \gamma k_j}\right] \frac{f(k_j)}{-\gamma c_2},\ 1 \le j \le N-1, \tag{9.85}$$

and $P(\gamma) = 0$ for $\gamma < \gamma_0$. We see by (9.85) that $P(\gamma)$ is discontinuous at the γ_j boundaries.

Let us now consider the optimal region boundaries $\gamma_0, \ldots, \gamma_{N-2}$. Solving (9.81) for $P_b(\gamma_j)$ yields

$$P_b(\gamma_j) = \bar{P}_b - \frac{1}{\lambda_1} - \frac{\lambda_2}{\lambda_1} \frac{P_{j+1}(\gamma_j) - P_j(\gamma_j)}{k_{j+1} - k_j},\ 0 \le j \le N-2, \tag{9.86}$$

where $k_0 = 0$ and $P_0(\gamma) = 0$. Unfortunately, this set of equations can be difficult to solve for the optimal boundary points $\{\gamma_j\}$. However, if we assume that $P(\gamma)$ is continuous at each boundary, then (9.86) becomes

$$P_b(\gamma_j) = \bar{P}_b - \frac{1}{\lambda_1},\ 0 \le j \le N-2, \tag{9.87}$$

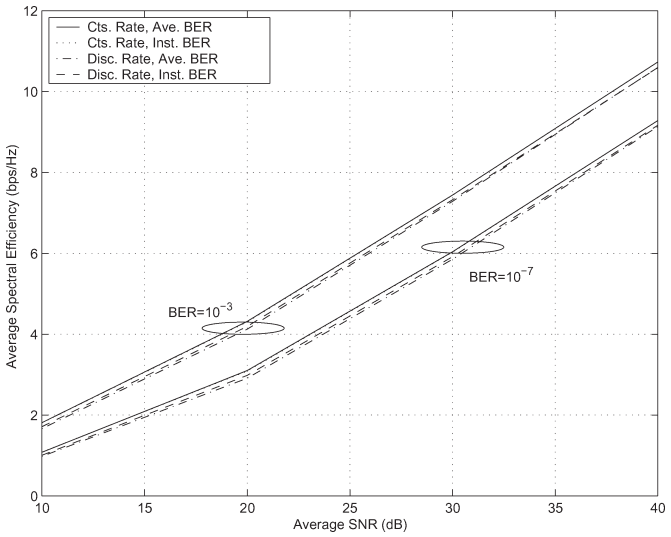


Figure 9.17: Spectral efficiency for different adaptation constraints.

for the Lagrangian λ_1. Under this assumption we can solve for the suboptimal rate region boundaries as

$$\gamma_{j-1} = \frac{f(k_j)}{k_j}\rho,\ 1 \le j \le N-1, \tag{9.88}$$

for some constant ρ. The constants λ_1 and ρ are found numerically such that the average power constraint

$$\sum_{j=1}^{N-1} \int_{\gamma_{j-1}}^{\gamma_j} \frac{P_j(\gamma)}{\bar{P}} p(\gamma)\, d\gamma = 1 \tag{9.89}$$

and the BER constraint (9.72) are satisfied. Note that the region boundaries (9.88) are suboptimal because $P(\gamma)$ is not necessarily continuous at the boundary regions, so these boundaries yield a suboptimal spectral efficiency.

In Figure 9.17 we plot average spectral efficiency for adaptive MQAM under both continuous and discrete rate adaptation, showing both average and instantaneous BER targets for a Rayleigh fading channel. The adaptive policies are based on the BER approximation (9.7) with a target BER of either 10^{-3} or 10^{-7}. For the discrete-rate cases we assume that six different MQAM signal constellations are available (seven fading regions), given by $\mathcal{M} = \{0, 4, 16, 64, 256, 1024, 4096\}$. We see in this figure that the spectral efficiencies of all four policies under the same instantaneous or average BER target are very close to each other. For discrete-rate adaptation, the spectral efficiency with an instantaneous BER target is slightly higher than with an average BER target even though the latter case is more constrained: that is because the efficiency with an average BER target is calculated using suboptimal rate region boundaries, which leads to a slight efficiency degradation.

9.5 Adaptive Techniques in Combined Fast and Slow Fading

In this section we examine adaptive techniques for composite fading channels consisting of both fast and slow fading (shadowing). We assume the fast fading changes too quickly to accurately measure and feed back to the transmitter, so the transmitter adapts only to the slow fading. The instantaneous SNR γ has distribution $p(\gamma|\bar{\gamma})$, where $\bar{\gamma}$ is a short-term average over the fast fading. This short-term average varies slowly because of shadowing and has a distribution $p(\bar{\gamma})$, where the average SNR relative to this distribution is $\bar{\bar{\gamma}}$. The transmitter adapts only to the slow fading $\bar{\gamma}$, hence its rate $k(\bar{\gamma})$ and power $P(\bar{\gamma})$ are functions of $\bar{\gamma}$. The power adaptation is subject to a long-term average power constraint over both the fast and slow fading:

$$\int_0^\infty P(\bar{\gamma}) p(\bar{\gamma})\, d\bar{\gamma} = \bar{P}. \tag{9.90}$$

As before, we approximate the instantaneous probability of bit error by the general form (9.44). Since the power and rate are functions of $\bar{\gamma}$, the conditional BER, conditioned on $\bar{\gamma}$, is

$$P_b(\gamma \mid \bar{\gamma}) \approx c_1 \exp\left[\frac{-c_2 \gamma P(\bar{\gamma})/\bar{P}}{2^{c_3 k(\bar{\gamma})} 3 - c_4}\right]. \tag{9.91}$$

Since the transmitter does not adapt to the fast fading γ, we cannot require a given instantaneous BER. However, since the transmitter adapts to the shadowing, we can require a target average probability of bit error averaged over the fast fading for a fixed value of the shadowing. This short-term average for a given $\bar{\gamma}$ is obtained by averaging $P_b(\gamma \mid \bar{\gamma})$ over the fast fading distribution $p(\gamma \mid \bar{\gamma})$:

$$\bar{P}_b(\bar{\gamma}) = \int_0^\infty P_b(\gamma \mid \bar{\gamma}) p(\gamma \mid \bar{\gamma})\, d\gamma. \tag{9.92}$$

Using (9.91) in (9.92) and assuming Rayleigh fading for the fast fading, this becomes

$$\bar{P}_b(\bar{\gamma}) = \frac{1}{\bar{\gamma}} \int_0^\infty c_1 \exp\left[\frac{-c_2 \gamma P(\bar{\gamma})/\bar{P}}{2^{c_3 k(\bar{\gamma})} - c_4} - \frac{\gamma}{\bar{\gamma}}\right] d\gamma = \frac{c_1}{\frac{c_2 \bar{\gamma} P(\bar{\gamma})/\bar{P}}{2^{c_3 k(\bar{\gamma})} - c_4} + 1}. \tag{9.93}$$

For example, with MQAM modulation with the tight BER bound (9.7), equation (9.93) becomes

$$\bar{P}_b(\bar{\gamma}) = \frac{.2}{\frac{c_2\bar{\gamma}P(\bar{\gamma})/\bar{P}}{2^{k(\bar{\gamma})}-1}+1}. \tag{9.94}$$

We can now invert (9.93) to obtain the adaptive rate $k(\bar{\gamma})$ as a function of the target average BER $\bar{P}_b$ and the power adaptation $P(\bar{\gamma})$:

$$k(\bar{\gamma}) = \frac{1}{c_3}\log_2\left(c_4 + \frac{K\bar{\gamma}P(\bar{\gamma})}{\bar{P}}\right), \tag{9.95}$$

where

$$K = \frac{c_2}{c_1/\bar{P}_b - 1} \tag{9.96}$$

depends only on the target average BER and decreases as this target decreases. We maximize spectral efficiency by maximizing

$$\mathbf{E}[k(\bar{\gamma})] = \int_0^\infty \frac{1}{c_3}\log_2\left(c_4 + \frac{K\bar{\gamma}P(\bar{\gamma})}{\bar{P}}\right)p(\bar{\gamma})\,d\bar{\gamma} \tag{9.97}$$

subject to the average power constraint (9.90).

Let us assume that $c_4 > 0$. Then this maximization and the power constraint are in the exact same form as (9.11) with the fading γ replaced by the slow fading $\bar{\gamma}$. Thus, the optimal power adaptation also has the same water-filling form as (9.13) and is given by

$$\frac{P(\bar{\gamma})}{\bar{P}} = \begin{cases} 1/\bar{\gamma}_0 - c_4/\bar{\gamma}K & \bar{\gamma} \geq c_4\bar{\gamma}_0/K, \\ 0 & \bar{\gamma} < c_4\bar{\gamma}_0/K, \end{cases} \tag{9.98}$$

where the channel is not used when $\bar{\gamma} < c_4\bar{\gamma}_0/K$. The value of $\bar{\gamma}_0$ is determined by the average power constraint. Substituting (9.98) into (9.95) yields the rate adaptation

$$k(\bar{\gamma}) = \frac{1}{c_3}\log_2\left(\frac{K\bar{\gamma}}{\bar{\gamma}_0}\right), \tag{9.99}$$

and the corresponding average spectral efficiency is given by

$$\frac{R}{B} = \int_{c_4\gamma_0/K}^\infty \log_2\left(\frac{K\bar{\gamma}}{\bar{\gamma}_0}\right)p(\bar{\gamma})\,d\bar{\gamma}. \tag{9.100}$$

Thus, we see that in a composite fading channel where rate and power are adapted only to the slow fading and for $c_4 > 0$ in (9.44), water-filling *relative to the slow fading* is the optimal power adaptation for maximizing spectral efficiency subject to an average BER constraint.

Our derivation has assumed that the fast fading is Rayleigh; however, it can be shown [30] that, with $c_4 > 0$ in (9.44), the optimal power and rate adaptation for any fast fading distribution have the same form. Since we have assumed $c_4 > 0$ in (9.44), the positivity constraint on power dictates the cutoff value below which the channel is not used. As we saw in Section 9.4.1, when $c_4 \leq 0$ the positivity constraint on rate dictates this cutoff, and the optimal power adaptation becomes inverse water-filling for $c_4 < 0$ and on–off power adaptation for $c_4 = 0$.

Chapter 9 Problems

1. Find the average SNR required to achieve an average BER of $\bar{P}_b = 10^{-3}$ for 8-PSK modulation in Rayleigh fading. What is the spectral efficiency of this scheme, assuming a symbol time of $T_s = 1/B$?

2. Consider a truncated channel inversion variable-power technique for Rayleigh fading with average SNR of 20 dB. What value of σ corresponds to an outage probability of .1? Find the maximum size MQAM constellation that can be transmitted under this policy so that, in non-outage, $P_b \approx 10^{-3}$.

3. Find the power adaptation for QPSK modulation that maintains a fixed $P_b = 10^{-3}$ in non-outage for a Rayleigh fading channel with $\bar{\gamma} = 20$ dB. What is the outage probability of this system?

4. Consider a variable-rate MQAM modulation scheme with just two constellations, $M = 4$ and $M = 16$. Assume a target P_b of approximately 10^{-3}. If the target cannot be met then no data is transmitted.

   (a) Using the BER bound (9.7), find the range of γ-values associated with the three possible transmission schemes (no transmission, 4-QAM, and 16-QAM) where the BER target is met. What is the cutoff γ_0 below which the channel is not used?

   (b) Assuming Rayleigh fading with $\bar{\gamma} = 20$ dB, find the average data rate of the variable-rate scheme.

   (c) Suppose that, instead of suspending transmission below γ_0, BPSK is transmitted for $0 \leq \gamma \leq \gamma_0$. Using the loose bound (9.6), find the average probability of error for this BPSK transmission.

5. Consider an adaptive modulation and coding scheme consisting of three modulations: BPSK, QPSK, and 8-PSK, along with three block codes of rate 1/2, 1/3, and 1/4. Assume that the first code provides roughly 3 dB of coding gain for each modulation type, the second code 4 dB, and the third code 5 dB. For each possible value of SNR $0 \leq \gamma \leq \infty$, find the combined coding and modulation with the maximum data rate for a target BER of 10^{-3} (you can use any reasonable approximation for modulation BER in this calculation, with SNR increased by the coding gain). Find the average data rate of the system for a Rayleigh fading channel with average SNR of 20 dB, assuming no transmission if the target BER cannot be met with any combination of modulation and coding.

6. Show that the spectral efficiency given by (9.11) with power constraint (9.12) is maximized by the water-filling power adaptation (9.13). Do this by setting up the Lagrangian equation, differentiating it, and solving for the maximizing power adaptation. Also show that, with this power adaptation, the rate adaptation is as given in (9.16).

7. In this problem we compare the spectral efficiency of nonadaptive techniques with that of adaptive techniques.

   (a) Using the tight BER bound for MQAM modulation given by (9.7), find an expression for the average probability of bit error in Rayleigh fading as a function of M and $\bar{\gamma}$.

   (b) Based on the expression found in part (a), find the maximum constellation size that can be transmitted over a Rayleigh fading channel with a target average BER of 10^{-3}, assuming $\bar{\gamma} = 20$ dB.

   (c) Compare the spectral efficiency of part (b) with that of adaptive modulation shown in Figure 9.3 for the same parameters. What is the spectral efficiency difference between the adaptive and nonadaptive techniques?

8. Consider a Rayleigh fading channel with an average SNR of 20 dB. Assume a target BER of 10^{-4}.

(a) Find the optimal rate and power adaptation for variable-rate variable-power MQAM as well as the cutoff value γ_0/K below which the channel is not used.

(b) Find the average spectral efficiency for the adaptive scheme derived in part (a).

(c) Compare your answer in part (b) to the spectral efficiency of truncated channel inversion, where γ_0 is chosen to maximize this efficiency.

9. Consider a discrete time-varying AWGN channel with four channel states. Assuming a fixed transmit power $\bar{P}$, the received SNR associated with each channel state is $\gamma_1 = 5$ dB, $\gamma_2 = 10$ dB, $\gamma_3 = 15$ dB, and $\gamma_4 = 20$ dB, respectively. The probabilities associated with the channel states are $p(\gamma_1) = .4$ and $p(\gamma_2) = p(\gamma_3) = p(\gamma_4) = .2$. Assume a target BER of 10^{-3}.

(a) Find the optimal power and rate adaptation for continuous-rate adaptive MQAM on this channel.

(b) Find the average spectral efficiency with this optimal adaptation.

(c) Find the truncated channel inversion power control policy for this channel and the maximum spectral efficiency that can be supported with this policy.

10. Consider a Rayleigh fading channel with an average received SNR of 20 dB and a required BER of 10^{-3}. Find the spectral efficiency of this channel using truncated channel inversion, assuming the constellation is restricted to size $0, 2, 4, 16, 64$, or 256.

11. Consider a Rayleigh fading channel with an average received SNR of 20 dB, a Doppler frequency of 80 Hz, and a required BER of 10^{-3}.

(a) Suppose you use adaptive MQAM modulation on this channel with constellations restricted to size $0, 2, 4, 16$, and 64. Using $\gamma_K^* = .1$, find the fading regions R_j associated with each of these constellations. Also find the average spectral efficiency of this restricted adaptive modulation scheme and the average time spent in each region R_j. If the symbol rate is $T_s = B^{-1}$, over approximately how many symbols is each constellation transmitted before a change in constellation size is needed?

(b) Find the exact BER of your adaptive scheme using (9.33). How does it differ from the target BER?

12. Consider a Rayleigh fading channel with an average received SNR of 20 dB, a signal bandwidth of 30 kHz, a Doppler frequency of 80 Hz, and a required BER of 10^{-3}.

(a) Assume the SNR estimate at the transmitter $\hat{\gamma}$ has the same distribution as the true channel SNR γ, so $p(\hat{\gamma}) = p(\gamma)$. For the optimal variable-rate variable-power MQAM scheme with no restrictions on rate or power, suppose that the transmit power and rate is adapted relative to $\hat{\gamma}$ instead of γ. Will the average transmitted data rate and average transmit power be larger, smaller, or the same as under perfect channel estimates ($\hat{\gamma} = \gamma$), and why? If over a given symbol time $\hat{\gamma} > \gamma$, will the probability of error associated with that symbol be larger or smaller than the target value and why?

(b) Suppose the estimation error $\varepsilon = \hat{\gamma}/\gamma$ in a variable-rate variable-power MQAM system with a target BER of 10^{-3} is uniformly distributed between .5 and 1.5. Find the resulting average probability of bit error for this system.

(c) Find an expression for the average probability of error in a variable-rate variable-power MQAM system in which the SNR estimate $\hat{\gamma}$ available at the transmitter is both a delayed and noisy estimate of γ : $\hat{\gamma}(t) = \gamma(t - \tau) + \gamma_\varepsilon(t)$. What joint distribution is needed to compute this average?

13. In this problem, we investigate the effect of channel estimation errors and delays in system performance. Consider a Rayleigh fading channel with mean power 1, uniform scattering and Doppler of f_D Hz. The channel output $y[t]$ is thus given by $y[t] = h[t]x[t] + \nu[t]$, where $h[t]$ is the fading realization, $x[t]$ is the transmitted symbol and $\nu[t]$ is the additive noise with power 1. Symbol time T_b s. We use adaptive modulation with no transmission, BPSK or 4-QAM; and the receiver feeds back the receiver output to the transmitter at the end of every symbol time. The transmitter uses $y[t]$ and $x[t]$ to estimate $\hat{h}[t+1]$ and uses this estimate to choose the constellation scheme to either have an outage or maintain a constant BER P_b for time $t+1$. The receiver knows the instantaneous channel state information $h[t]$ and uses that for decoding.

    (a) Find the two threshold received SNRs $\gamma_0, \gamma_1 (\gamma_0 < \gamma_1)$ at which the transmitter switches modulations. Use the exact AWGN error probability expressions wherever possible.

    (b) Find out the MMSE (minimum mean squared error) estimate of $h[t]$ given $y[t]$ and $x[t]$. This is the effect of the channel estimation error.

    (c) Find out the MMSE estimate of $h[t+1]$ given the estimate of $h[t]$. This is the effect of delay in feedback.

    (d) Assuming that the transmitter uses the MMSE estimate $\hat{h}[t+1]$ to choose the transmit modulation for time $t+1$, and the actual realization is $h[t+1]$ known to the receiver, write down an expression for the bit error rate $P_b(|h[t+1]|, |\hat{h}[t+1]|)$. Further, write the expression for the BER estimated over the joint distribution of $y[t]$ and $h[t+1]$.

    Note: In general the optimal transmitter would want to use past values of the feedback to refine its estimate of $h[t+1]$ but in this problem we do not consider that possibility.

14. Consider an adaptive trellis-coded MQAM system with a coding gain of 3 dB. Assume a Rayleigh fading channel with an average received SNR of 20 dB. Find the optimal adaptive power and rate policy for this system and the corresponding average spectral efficiency. Assume a target BER of 10^{-3}.

15. In Chapter 6, a bound on P_b for nonrectangular MQAM was given as

$$P_b \approx \frac{4}{\log_2 M} Q\left(\sqrt{\frac{3\gamma}{(M-1)}}\right).$$

    Find values for c_1, c_2, c_3, and c_4 for the general BER form (9.44) to approximate this bound with $M = 8$. Any curve approximation technique is acceptable. Plot both BER formulas for $0 \leq \gamma \leq 30$ dB.

16. Show that the average spectral efficiency $\mathbf{E}[k(\gamma)]$ for $k(\gamma)$ given by (9.45) with power constraint $\bar{P}$ is maximized by the power adaptation (9.48).

17. In this problem we investigate the optimal adaptive modulation for MPSK modulation based on the three BER bounds (9.52), (9.53), and (9.54). We assume a Rayleigh fading channel (so that γ is exponentially distributed with $\bar{\gamma} = 30$ dB) and a target BER of $P_b = 10^{-7}$.

    (a) The cutoff fade depth γ_0 must satisfy

$$\int_{\gamma_0/K}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma K}\right) p(\gamma)\, d\gamma \leq 1$$

    for K given by (9.49). Find the cutoff value γ_0 corresponding to the power adaptation for each of the three bounds.

(b) Plot $P(\gamma)/\bar{P}$ and $k(\gamma)$ as a function of γ for Bounds 1, 2, and 3 for γ ranging from 0 dB to 30 dB. Also state whether the cutoff value below which the channel is not used is based on the power or rate positivity constraint.

(c) How does the power adaptation associated with the different bounds differ at low SNRs? What about at high SNRs?

18. Show that, under discrete rate adaptation for general M-ary modulation, the power adaptation that maintains a target instantaneous BER is given by (9.64). Also show that the region boundaries that maximize spectral efficiency–obtained using the Lagrangian given in (9.66) – are given by (9.68) and (9.69).

19. Show that, for general M-ary modulation with an average target BER, the Lagrangian (9.80) implies that the optimal power and BER adaptation must satisfy (9.82). Then show how (9.82) leads to BER adaptation given by (9.83), which in turn leads to the power adaptation given by (9.84) and (9.85). Finally, use (9.81) to show that the optimal rate region boundaries must satisfy (9.86).

20. Consider adaptive MPSK where the constellation is restricted to either no transmission or $M = 2, 4, 8, 16$. Assume the probability of error is approximated using (9.52). Find and plot the optimal discrete-rate continuous-power adaptation for $0 \leq \gamma \leq 30$ dB assuming a Rayleigh fading channel with $\bar{\gamma} = 20$ dB and a target P_b of 10^{-4}. What is the resulting average spectral efficiency?

21. We assume the same discrete-rate adaptive MPSK as in the previous problem, except now there is an average target P_b of 10^{-4} instead of an instantaneous target. Find the optimal discrete-rate continuous-power adaptation for a Rayleigh fading channel with $\bar{\gamma} = 20$ dB and the corresponding average spectral efficiency.

22. Consider a composite fading channel with fast Rayleigh fading and slow log-normal shadowing with an average dB SNR $\mu_{\psi_{\rm dB}} = 20$ dB (averaged over both fast and slow fading) and $\sigma_{\psi_{\rm dB}} = 8$ dB. Assume an adaptive MPSK modulation that adapts only to the shadowing, with a target average BER of 10^{-3}. Using the BER approximation (9.52), find the optimal power and rate adaptation policies as a function of the slow fading $\bar{\gamma}$ that maximize average spectral efficiency while meeting the average BER target. Also determine the average spectral efficiency that results from these policies.

23. In Section 9.5 we determined the optimal adaptive rate and power policies to maximize average spectral efficiency while meeting a target average BER in combined Rayleigh fading and shadowing. The derivation assumed the general bound (9.44) with $c_4 > 0$. For the same composite channel, find the optimal adaptive rate and power policies to maximize average spectral efficiency while meeting a target average BER assuming $c_4 < 0$. *Hint:* The derivation is similar to the case of continuous-rate adaptation using the second MPSK bound and results in the same channel inversion power control.

24. As in the previous problem, we again examine the adaptive rate and power policies to maximize average spectral efficiency while meeting a target average BER in combined Rayleigh fading and shadowing. In this problem we assume the general bound (9.44) with $c_4 = 0$. For the composite channel, find the optimal adaptive rate and power policies to maximize average spectral efficiency while meeting a target average BER assuming $c_4 = 0$. *Hint*: the derivation is similar to that of Section 9.4.1 for the third MPSK bound and results in the same on-off power control.

Bibliography

[1] J. F. Hayes, "Adaptive feedback communications," *IEEE Trans. Commun. Tech.*, pp. 29–34, February 1968.

[2] J. K. Cavers, "Variable-rate transmission for Rayleigh fading channels," *IEEE Trans. Commun.*, pp. 15–22, February 1972.

[3] S. Otsuki, S. Sampei, and N. Morinaga, "Square-QAM adaptive modulation/TDMA/TDD systems using modulation level estimation with Walsh function," *Elec. Lett.*, pp. 169–71, February 1995.

[4] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, pp. 2223–30, July 1995.

[5] Y. Kamio, S. Sampei, H. Sasaoka, and N. Morinaga, "Performance of modulation-level-controlled adaptive-modulation under limited transmission delay time for land mobile communications," *Proc. IEEE Veh. Tech. Conf.*, pp. 221–5, July 1995.

[6] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Commun.*, pp. 653–63, May 1991.

[7] M. Rice and S. B. Wicker, "Adaptive error control for slowly varying channels," *IEEE Trans. Commun.*, pp. 917–26, February–April 1994.

[8] S. M. Alamouti and S. Kallel, "Adaptive trellis-coded multiple-phased-shift keying for Rayleigh fading channels," *IEEE Trans. Commun.*, pp. 2305–14, June 1994.

[9] T. Ue, S. Sampei, and N. Morinaga, "Symbol rate and modulation level controlled adaptive modulation/TDMA/TDD for personal communication systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 306–10, July 1995.

[10] H. Matsuoka, and Y. Kamio, "Symbol rate and modulation level controlled adaptive modulation/TDMA/TDD for personal communication systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 487–91, April 1996.

[11] S. Sampei, N. Morinaga, and Y. Kamio, "Adaptive modulation/TDMA with a BDDFE for 2 Mbit/s multimedia wireless communication systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 311–15, July 1995.

[12] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[13] K. M. Kamath and D. L. Goeckel, "Adaptive-modulation schemes for minimum outage probability in wireless systems," *IEEE Trans. Commun.*, pp. 1632–5, October 2004.

[14] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, pp. 389–400, April 1988.

[15] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *Bell System Tech. J.,* pp. 429–56, February 1983.

[16] G. D. Forney, Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Quereshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun*., pp. 632–47, September 1984.

[17] J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001.

[18] M. Filip and E. Vilar, "Implementation of adaptive modulation as a fade countermeasure," *Internat. J. Sat. Commun*., pp. 181–91, 1994.

[19] C. C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun*., pp. 2032–40, December 2000.

[20] L. Kleinrock, *Queueing Systems*, vol. I: *Theory*, Wiley, New York, 1975.

[21] W. C. Jakes, Jr., *Microwave Mobile Communications*, Wiley, New York, 1974.

[22] H. S. Wang and P.-C. Chang, "On verifying the first-order Markov assumption for a Rayleigh fading channel model," *IEEE Trans.Veh.Tech*., pp. 353–7, May 1996.

[23] X. Tang, M.-S. Alouini, and A. Goldsmith, " Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Trans. Commun*., pp. 1856–64, December 1999.

[24] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Tech*., pp. 686–93, November 1991.

[25] A. J. Goldsmith and L. J. Greenstein, "Effect of average power estimation error on adaptive MQAM modulation," *Proc. IEEE Internat. Conf. Commun*., pp. 1105–9, June 1997.

[26] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun*., pp. 844–55, June 1999.

[27] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Proc. Mag*., pp. 62–75, May 2000.

[28] M.-S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer J. Wireless Pers. Commun*., pp. 119–43, May 2000.

[29] S.-G. Chua and A. J. Goldsmith, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun*., pp. 595–602, May 1998.

[30] S. Vishwanath, S. A. Jafar, and A. J. Goldsmith, "Adaptive resource allocation in composite fading environments," *Proc. IEEE Globecom Conf*., pp. 1312–16, November 2001.

Chapter 10

MIMO Communications

In this chapter we consider systems with multiple antennas at the transmitter and receiver, which are commonly referred to as multiple-input multiple-output (MIMO) systems. The multiple antennas can be used to increase data rates through multiplexing or to improve performance through diversity. We have already seen diversity in Chapter 7. In MIMO systems, the transmit and receive antennas can both be used for diversity gain. Multiplexing exploits the structure of the channel gain matrix to obtain independent signaling paths that can be used to send independent data. Indeed, the initial excitement about MIMO was sparked by the pioneering work of Winters [1], Paulraj and Kailath [2], Foschini [3], Foschini and Gans [4], and Telatar [5] predicting remarkable spectral efficiencies for wireless systems with multiple transmit and receive antennas. These spectral efficiency gains often require accurate knowledge of the channel at the receiver – and sometimes at the transmitter as well. In addition to spectral efficiency gains, ISI and interference from other users can be reduced using smart antenna techniques. The cost of the performance enhancements obtained through MIMO techniques is the added cost of deploying multiple antennas, the space and circuit power requirements of these extra antennas (especially on small handheld units), and the added complexity required for multidimensional signal processing. In this chapter we examine the different uses for multiple antennas and find their performance advantages. This chapter uses several key results from matrix theory: Appendix C provides a brief overview of these results. Comprehensive books on MIMO systems include [6, 7, 8, 9].

10.1 Narrowband MIMO Model

In this section we consider a narrowband MIMO channel such that the signal bandwidth is much less than the coherence bandwidth of the channel between each transmit/receive antenna pair. Thus, ISI can be neglected and there is flat-fading on the channel between each transmit and receive antenna. A narrowband point-to-point communication system of M_t transmit and M_r receive antennas is shown in Figure 10.1. This system can be represented by the following discrete-time complex baseband model:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{M_r} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{1M_t} \\ \vdots & \ddots & \vdots \\ h_{M_r1} & \cdots & h_{M_rM_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{M_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_{M_r} \end{bmatrix} \tag{10.1}$$

or simply as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{10.2}$$

where $\mathbf{x}$ represents the M_t-dimensional transmitted symbol taken from an alphabet $\mathcal{X}^{M_t}$, $\mathbf{n}$ is the M_r-dimensional complex noise vector, and $\mathbf{H}$ is the $M_r \times M_t$ matrix of complex channel gains h_{ij} representing the gain from transmit antenna j to receive antenna i. We assume a channel bandwidth of B and complex Gaussian noise with zero mean and covariance matrix $\sigma^2 \mathbf{I}_{M_r}$, where typically $\sigma^2 \triangleq \mathbf{E}[n_i^2] = N_0 B$, the noise power in the ith receive antenna. For simplicity, given a transmit power constraint P we will assume an equivalent model with a noise power σ^2 of unity and transmit power $P/\sigma^2 = \rho$, where ρ can be interpreted as the average SNR per receive antenna under unity channel gain. This power constraint implies that the input symbols satisfy

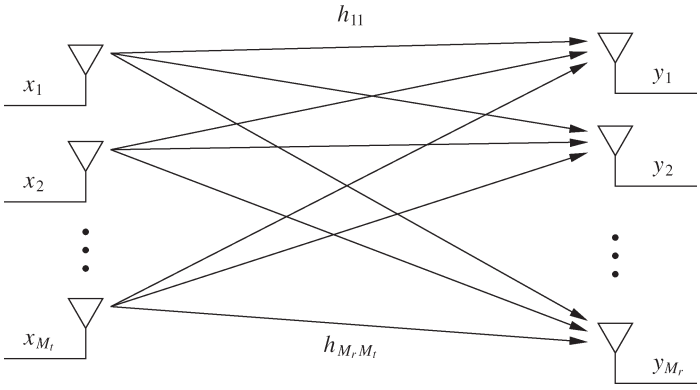


Figure 10.1: MIMO systems.

$$\sum_{i=1}^{M_t} \mathbf{E}[x_i x_i^*] = \rho, \tag{10.3}$$

or (equivalently) that $\mathrm{Tr}(\mathbf{R_x}) = \rho$, where $\mathrm{Tr}(\mathbf{R_x})$ is the trace of the input covariance matrix $\mathbf{R_x} = \mathbf{E}[\mathbf{x}\mathbf{x}^H]$.

Different assumptions can be made about the knowledge of the channel gain matrix $\mathbf{H}$ at the transmitter and receiver, referred to as channel side information at the transmitter (CSIT) and channel side information at the receiver (CSIR), respectively. For a static channel CSIR is typically assumed, since the channel gains can be obtained fairly easily by sending a pilot sequence for channel estimation. More details on estimation techniques for MIMO channels can be found in [9, Chap. 3.9]. If a feedback path is available then CSIR from the receiver can be sent back to the transmitter to provide CSIT: CSIT may also be available in bidirectional systems without a feedback path when the reciprocal properties of propagation are exploited. When the channel is not known to either the transmitter or receiver then some distribution on the channel gain matrix must be assumed. The most common model for this distribution is a zero-mean spatially white (ZMSW) model, where the entries of $\mathbf{H}$ are assumed to be independent and identically distributed (i.i.d.) zero-mean, unit-variance, complex circularly symmetric Gaussian random variables.[1] We adopt this model unless stated otherwise. Alternatively, these entries may be complex circularly symmetric Gaussian random variables with a nonzero mean or with a covariance matrix not equal to the identity matrix. In general, different assumptions about CSI and about the distribution of the $\mathbf{H}$ entries lead to different channel capacities and different approaches to space-time signaling.

Optimal decoding of the received signal requires maximum likelihood demodulation. If the symbols modulated onto each of the M_t transmit antennas are chosen from an alphabet of size $|\mathcal{X}|$, then – because of the

[1] A complex random vector $\mathbf{x}$ is circularly symmetric if, for any $\theta \in [0, 2\pi]$, the distribution of $\mathbf{x}$ is the same as the distribution of $e^{j\theta}\mathbf{x}$. Thus, the real and imaginary parts of $\mathbf{x}$ are i.i.d. For $\mathbf{x}$ circularly symmetric, taking $\theta = \pi$ implies that $\mathbf{E}[\mathbf{x}] = 0$. Similarly, taking $\theta = \pi/2$ implies that $\mathbf{E}[\mathbf{x}\mathbf{x}^T] = 0$, which by definition means that $\mathbf{x}$ is a *proper* random vector. For a complex random vector $\mathbf{x}$ the converse is also true: if $\mathbf{x}$ is zero-mean and proper then it is circularly symmetric.

cross-coupling between transmitted symbols at the receiver antennas – ML demodulation requires an exhaustive search over all $|\mathcal{X}|^{M_t}$ possible input vectors of M_t symbols. For general channel matrices, when the transmitter does not know $\mathbf{H}$, the complexity cannot be reduced further. This decoding complexity is typically prohibitive for even a small number of transmit antennas. However, decoding complexity is significantly reduced if the channel is known to the transmitter, as shown in the next section.

10.2 Spatial Dimensions of the MIMO Channel

We have seen in Chapter 7 that multiple antennas at the transmitter or receiver can be used for diversity gain. When both the transmitter and receiver have multiple antennas, there is another mechanism for performance gain called *multiplexing gain*. The multiplexing gain of a MIMO system results from the fact that a MIMO channel can be decomposed into a number R of parallel independent channels. By multiplexing independent data onto these independent channels, we get an R-fold increase in data rate in comparison to a system with just one antenna at the transmitter and receiver. This increased data rate is called the multiplexing gain. In this section we describe how to obtain independent channels from a MIMO system by using transmit precoding and receiver combining based on the channel's singular value decomposition.

10.2.1 Singular Value Decomposition of the Channel Matrix

Consider a MIMO channel with $M_r \times M_t$ channel gain matrix $\mathbf{H}$ that is known to both the transmitter and the receiver. Let $R_{\mathbf{H}}$ denote the rank of $\mathbf{H}$. From Appendix C, for any matrix $\mathbf{H}$ we can obtain its singular value decomposition (SVD) as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{10.4}$$

where the $M_r \times M_r$ matrix $\mathbf{U}$ and the $M_t \times M_t$ matrix $\mathbf{V}$ are unitary matrices[2] and where $\mathbf{\Sigma}$ is an $M_r \times M_t$ diagonal matrix of singular values $\{\sigma_i\}$ of $\mathbf{H}$. These singular values have the property that $\sigma_i = \sqrt{\lambda_i}$ for λ_i the ith largest eigenvalue of $\mathbf{H}\mathbf{H}^H$, and $R_{\mathbf{H}}$ of these singular values are nonzero. Because $R_{\mathbf{H}}$, the rank of matrix $\mathbf{H}$, cannot exceed the number of columns or rows of $\mathbf{H}$, it follows that $R_{\mathbf{H}} \leq \min(M_t, M_r)$. If $\mathbf{H}$ is full rank, which is referred to as a *rich scattering environment*, then $R_{\mathbf{H}} = \min(M_t, M_r)$. Other environments may lead to a low-rank $\mathbf{H}$: a channel with high correlation among the gains in $\mathbf{H}$ may have rank 1.

10.2.2 Transmit Precoding and Receiver Combining

The parallel decomposition of the channel is obtained by defining linear transformations on the channel input and output $\mathbf{x}$ and $\mathbf{y}$ via transmitter *precoding* and receiver *combining*.[3] In transmitter precoding the input $\mathbf{x}$ to the antennas is generated by a linear transformation on input vector $\tilde{\mathbf{x}}$ as $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$. Receiver combining performs a similar operation at the receiver by multiplying the channel output $\mathbf{y}$ by $\mathbf{U}^H$, as shown in Figure 10.2.

The transmitter precoding and receiver combining transform the MIMO channel into $R_{\mathbf{H}}$ parallel single-input single-output (SISO) channels with input $\tilde{\mathbf{x}}$ and output $\tilde{\mathbf{y}}$, since from the SVD we have that

[2] $\mathbf{U}$ and $\mathbf{V}$ unitary imply that $\mathbf{U}^H\mathbf{U} = \mathbf{I}_{M_r}$ and $\mathbf{V}^H\mathbf{V} = \mathbf{I}_{M_t}$.

[3] The terminology of precoding and combining in the context of the MIMO channel decomposition is somewhat misleading, as the precoder operates on coded bits, and the combiner does not combine the signals received over all antennas into a single spatial stream, as is the case with diversity combining. Nevertheless, this terminology has been widely adopted and hence we use it here

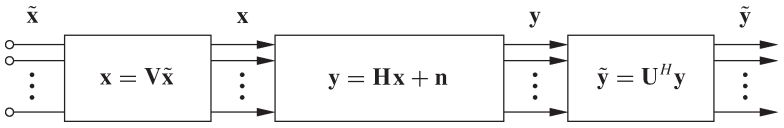


Figure 10.2: Transmitter precoding and receiver combining.

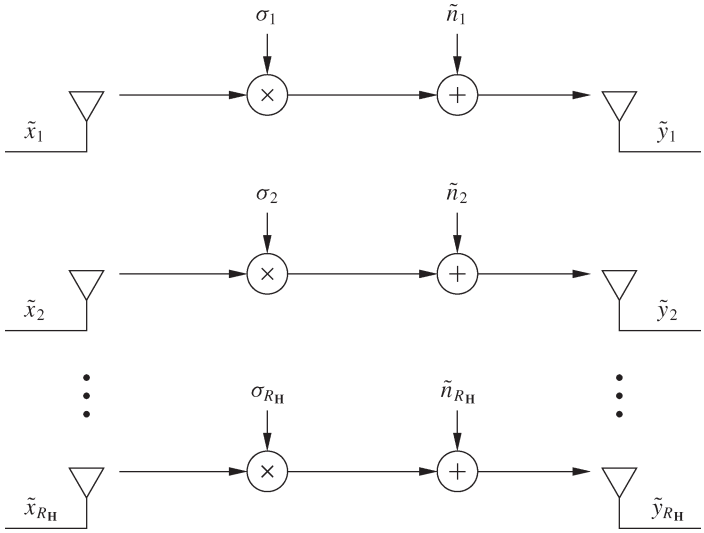


Figure 10.3: Parallel decomposition of the MIMO channel.

$$\begin{aligned}
\tilde{\mathbf{y}} &= \mathbf{U}^H(\mathbf{H}\mathbf{x} + \mathbf{n}) \\
&= \mathbf{U}^H(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{x} + \mathbf{n}) \\
&= \mathbf{U}^H(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}} + \mathbf{n}) \\
&= \mathbf{U}^H\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}} + \mathbf{U}^H\mathbf{n} \\
&= \mathbf{\Sigma}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}, \qquad (10.5)
\end{aligned}$$

where $\tilde{\mathbf{n}} = \mathbf{U}^H\mathbf{n}$ and where $\mathbf{\Sigma}$ is the matrix of singular values of $\mathbf{H}$ with σ_i on the ith diagonal and zeros everywhere else. Note that multiplication by a unitary matrix does not change the distribution of the noise – that is, $\mathbf{n}$ and $\tilde{\mathbf{n}}$ are identically distributed. Thus, the transmitter precoding and receiver combining transform the MIMO channel into $R_\mathbf{H}$ parallel independent channels, where the ith channel has input $\tilde{x}_i$, output $\tilde{y}_i$, noise $\tilde{n}_i$, and channel gain σ_i. Note that the σ_i are related because they are all functions of $\mathbf{H}$, but since the resulting parallel channels do not interfere with each other we say that the channels with these gains are independent – linked only through the total power constraint. This parallel decomposition is shown in Figure 10.3. Since the parallel channels do not interfere with each other, the optimal ML demodulation complexity is linear in $R_\mathbf{H}$, the number of independent paths that need to be demodulated. Moreover, by sending independent data across each of the parallel channels, the MIMO channel can support $R_\mathbf{H}$ times the data rate of a system with just one transmit and receive antenna, leading to a multiplexing gain of $R_\mathbf{H}$. Note, however, that the performance on each one of the channels will depend on its gain σ_i. The next section will more precisely characterize the multiplexing gain associated with the Shannon capacity of the MIMO channel.

Example 10.1: Find the equivalent parallel channel model for a MIMO channel with channel gain matrix

$$\mathbf{H} = \begin{bmatrix} .1 & .3 & .7 \\ .5 & .4 & .1 \\ .2 & .6 & .8 \end{bmatrix}. \tag{10.6}$$

Solution: The SVD of $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ is given by

$$\mathbf{H} = \begin{bmatrix} -.555 & .3764 & -.7418 \\ -.3338 & -.9176 & -.2158 \\ -.7619 & .1278 & .6349 \end{bmatrix} \begin{bmatrix} 1.3333 & 0 & 0 \\ 0 & .5129 & 0 \\ 0 & 0 & .0965 \end{bmatrix} \times \begin{bmatrix} -.2811 & -.7713 & -.5710 \\ -.5679 & -.3459 & .7469 \\ -.7736 & .5342 & -.3408 \end{bmatrix}. \tag{10.7}$$

There are three nonzero singular values and so $R_{\mathbf{H}} = 3$, leading to three parallel channels with respective channel gains $\sigma_1 = 1.3333, \sigma_2 = .5129$, and $\sigma_3 = .0965$. Note that the channels have diminishing gain, with a very small gain on the third channel. Hence, this last channel will either have a high error probability or a low capacity.

The parallel decomposition of the channel $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ requires perfect CSIT for transmitter precoding (multiplication of the input vector by $\mathbf{V}$) and perfect CSIR for receiver combining (multiplication of the output vector by $\mathbf{U}^H$). In practice the channel matrix's $M_r \times M_t$ values are typically estimated at the receiver to obtain CSIR and this estimate is fed back to the transmitter to obtain CSIT. For rapidly-varying channels, the coherence time may be too short for the system to accurately estimate the channel matrix at the receiver, let alone feed this estimate back to the transmitter before the channel changes. Moreover, the limited capacity of the feedback channel may require low-rate quantization of the channel matrix estimate, which introduces quantization noise. These practical limitations result in CSIT and CSIR that may differ significantly from the actual channel $\mathbf{H}$. When the channel decomposition is based on incorrect estimates of the channel, the spatial streams associated with the decomposition are no longer independent and hence interfere with each other. To mitigate this effect, precoding and combining techniques that take into account imperfect channel estimates have been developed. Discussion of these techniques can be found in [6, Chap. 5.10],[10, 11].

10.3 MIMO Channel Capacity

This section focuses on the Shannon capacity of a MIMO channel, which equals the maximum data rate that can be transmitted over the channel with arbitrarily small error probability. Capacity versus outage defines the maximum rate that can be transmitted over the channel with some nonzero outage probability. Channel capacity depends on what is known about the channel gain matrix or its distribution at the transmitter and/or receiver. First the static channel capacity under different assumptions about this channel knowledge will be given, which forms the basis for the subsequent section on capacity of fading channels.

10.3.1 Static Channels

The capacity of a MIMO channel is an extension of the mutual information formula for a SISO channel given by equation (4.6) to a matrix channel. For static channels a good estimate of $\mathbf{H}$ can be obtained fairly easily at the

receiver, so we assume CSIR throughout this section. Under this assumption, the capacity is given in terms of the mutual information between the channel input vector $\mathbf{x}$ and output vector $\mathbf{y}$ as

$$C = \max_{p(\mathbf{x})} I(\mathbf{X};\mathbf{Y}) = \max_{p(\mathbf{x})} [H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X})] \tag{10.8}$$

for $H(\mathbf{Y})$ and $H(\mathbf{Y}|\mathbf{X})$ the entropy in $\mathbf{y}$ and $\mathbf{y}|\mathbf{x}$, as defined in Section 4.1[4]. The definition of entropy yields that $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{n})$, the entropy in the noise. Since this noise $\mathbf{n}$ has fixed entropy independent of the channel input, maximizing mutual information is equivalent to maximizing the entropy in $\mathbf{y}$.

Given covariance matrix $\mathbf{R_x}$ on the input vector $\mathbf{x}$, the output covariance matrix $\mathbf{R_y}$ associated with MIMO channel output $\mathbf{y}$ is given by

$$\mathbf{R_y} = \mathbf{E}\left[\mathbf{y}\mathbf{y}^H\right] = \mathbf{H}\mathbf{R_x}\mathbf{H}^H + \mathbf{I}_{M_r}. \tag{10.9}$$

It turns out that, for all random vectors with a given covariance matrix $\mathbf{R_y}$, the entropy of $\mathbf{y}$ is maximized when $\mathbf{y}$ is a zero-mean, circularly symmetric complex Gaussian (ZMCSCG) random vector [5]. But $\mathbf{y}$ is ZMCSCG only if the input $\mathbf{x}$ is ZMCSCG, and hence this is the optimal distribution on $\mathbf{x}$ in (10.8), subject to the power constraint $\mathrm{Tr}(\mathbf{R_x}) = \rho$. Recall that $\det[\mathbf{A}]$ denotes the determinant of a square matrix $\mathbf{A}$. Thus we have $H(\mathbf{Y}) = B\log_2\det[\pi e\mathbf{R_y}]$ and $H(\mathbf{n}) = B\log_2\det[\pi e\mathbf{I}_{M_r}]$, resulting in the mutual information

$$I(\mathbf{X};\mathbf{Y}) = B\log_2\det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H]. \tag{10.10}$$

This formula was derived in [4, 5] for the mutual information of a multiantenna system, and it also appeared in earlier works on MIMO systems [12, 13] and matrix models for ISI channels [14, 15].

The MIMO capacity is achieved by maximizing the mutual information (10.10) over all input covariance matrices $\mathbf{R_x}$ satisfying the power constraint:

$$C = \max_{\mathbf{R_x}:\mathrm{Tr}(\mathbf{R_x})=\rho} B\log_2\det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H]. \tag{10.11}$$

Clearly the optimization relative to $\mathbf{R_x}$ will depend on whether or not $\mathbf{H}$ is known at the transmitter. We now consider this optimization under different assumptions about transmitter CSI.

CHANNEL KNOWN AT TRANSMITTER: WATER-FILLING

The MIMO decomposition described in Section 10.2 allows a simple characterization of the MIMO channel capacity for a fixed channel matrix $\mathbf{H}$ known at the transmitter and receiver. Specifically, the capacity equals the sum of capacities on each of the independent parallel channels with the transmit power optimally allocated between these channels. This optimization of transmit power across the independent channels results from optimizing the input covariance matrix to maximize the capacity formula (10.11). Substituting the matrix SVD (10.4) into (10.11) and using properties of unitary matrices, we get the MIMO capacity with CSIT and CSIR as

$$C = \max_{\rho_i:\Sigma_i\rho_i\le\rho} \sum_{i=1}^{R_\mathbf{H}} B\log_2(1+\sigma_i^2\rho_i), \tag{10.12}$$

where $R_\mathbf{H}$ is the number of nonzero singular values σ_i^2 of $\mathbf{H}$. Since the MIMO channel decomposes into $R_\mathbf{H}$ parallel channels, we say that it has $R_\mathbf{H}$ *degrees of freedom*. Since $\rho = P/\sigma^2$, the capacity (10.12) can also be expressed in terms of the power allocation P_i to the ith parallel channel as

[4]Entropy was defined in Section 4.1 for scalar random variables, and the definition is identical for random vectors.

$$C = \max_{P_i:\Sigma_i P_i \le P} \sum_{i=1}^{R_{\mathbf{H}}} B \log_2 \left(1 + \frac{\sigma_i^2 P_i}{\sigma^2}\right) = \max_{P_i:\Sigma_i P_i \le P} \sum_{i=1}^{R_{\mathbf{H}}} B \log_2 \left(1 + \frac{P_i \gamma_i}{P}\right), \tag{10.13}$$

where $\gamma_i = \sigma_i^2 P/\sigma^2$ is the SNR associated with the ith channel at full power. This expression indicates that, at high SNRs, channel capacity increases linearly with the number of degrees of freedom in the channel. Conversely, at low SNRs, all power will be allocated to the parallel channel with the largest SNR (or, equivalently, the largest σ_i^2). The capacity formula (10.13) is similar to the case of flat fading (4.13) or frequency-selective fading (4.31). Solving the optimization leads to a water-filling power allocation for the MIMO channel:

$$\frac{P_i}{P} = \begin{cases} 1/\gamma_0 - 1/\gamma_i & \gamma_i \ge \gamma_0, \\ 0 & \gamma_i < \gamma_0, \end{cases} \tag{10.14}$$

for some cutoff value γ_0. The resulting capacity is then

$$C = \sum_{i:\gamma_i \ge \gamma_0} B \log_2 \left(\frac{\gamma_i}{\gamma_0}\right). \tag{10.15}$$

Example 10.2: Find the capacity and optimal power allocation for the MIMO channel given in Example 10.1, assuming $\rho = P/\sigma^2 = 10$ dB and $B = 1$ Hz.

Solution: From Example 10.2.2, the singular values of the channel are $\sigma_1 = 1.3333, \sigma_2 = 0.5129$, and $\sigma_3 = 0.0965$. Since $\gamma_i = 10\sigma_i^2$, this means we have $\gamma_1 = 17.7769, \gamma_2 = 2.6307$, and $\gamma_3 = .0931$. Assuming that power is allocated to all three parallel channels, the power constraint yields

$$\sum_{i=1}^{3} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_i}\right) = 1 \Longrightarrow \frac{3}{\gamma_0} = 1 + \sum_{i=1}^{3} \frac{1}{\gamma_i} = 12.1749. \tag{10.16}$$

Solving for γ_0 yields $\gamma_0 = .2685$, which is inconsistent because $\gamma_3 = .0931 < \gamma_0 = .2685$. Thus, the third channel is not allocated any power. Then the power constraint yields

$$\sum_{i=1}^{2} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_i}\right) = 1 \Longrightarrow \frac{2}{\gamma_0} = 1 + \sum_{i=1}^{2} \frac{1}{\gamma_i} = 1.4364. \tag{10.17}$$

Solving for γ_0 in this case yields $\gamma_0 = 1.392 < \gamma_2$, so this is the correct cutoff value. Then $P_i/P = 1/1.392 - 1/\gamma_i$, so $P_1/P = .662$ and $P_2/P = .338$. The capacity is given by $C = B \log_2(\gamma_1/\gamma_0) + B \log_2(\gamma_2/\gamma_0) = 4.59$.

Capacity under perfect CSIT and CSIR can also be defined on channels for which there is a single antenna at the transmitter and multiple receive antennas (single-input multiple-output, SIMO) or multiple transmit antennas and a single receive antenna (multiple-input single-output, MISO). These channels can obtain diversity and array gain from the multiple antennas, but no multiplexing gain. If both transmitter and receiver know the channel then the capacity equals that of an SISO channel with the signal transmitted or received over the multiple antennas coherently combined to maximize the channel SNR, as in MRC. This results in capacity $C = B \log_2(1 + \rho \|\mathbf{h}\|^2)$ where the channel matrix $\mathbf{H}$ is reduced to a vector $\mathbf{h}$ of channel gains, the optimal weight vector $\mathbf{c} = \mathbf{h}^H/\|\mathbf{h}\|$, and $\rho = P/\sigma^2$.

CHANNEL UNKNOWN AT TRANSMITTER: UNIFORM POWER ALLOCATION

Suppose now that the receiver knows the channel but the transmitter does not. Without channel information, the transmitter cannot optimize its power allocation or input covariance structure across antennas. If the distribution of $\mathbf{H}$ follows the ZMSW channel gain model, then there is no bias in terms of the mean or covariance of $\mathbf{H}$. Thus, it seems intuitive that the best strategy should be to allocate equal power to each transmit antenna, resulting in an input covariance matrix equal to the scaled identity matrix: $\mathbf{R}_{\mathbf{x}} = (\rho/M_t)\mathbf{I}_{M_t}$. It is shown in [5] that, under these assumptions, this input covariance matrix indeed maximizes the mutual information of the channel. For an M_t-transmit M_r-receive antenna system, this yields mutual information given by

$$I(\mathbf{x};\mathbf{y}) = B\log_2 \det\left[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H\right]. \tag{10.18}$$

Using the SVD of $\mathbf{H}$, we can express this as

$$I(\mathbf{x};\mathbf{y}) = \sum_{i=1}^{R_{\mathbf{H}}} B\log_2\left(1 + \frac{\gamma_i}{M_t}\right), \tag{10.19}$$

where $\gamma_i = \sigma_i^2\rho = \sigma_i^2 P/\sigma^2$.

The mutual information of the MIMO channel (10.19) depends on the specific realization of the matrix $\mathbf{H}$, in particular its singular values $\{\sigma_i\}$. The average mutual information of a random matrix $\mathbf{H}$, averaged over the matrix distribution, depends on the probability distribution of the singular values of $\mathbf{H}$ [5, 16, 17]. In fading channels the transmitter can transmit at a rate equal to this average mutual information and ensure correct reception of the data, as discussed in the next section. But for a static channel, if the transmitter does not know the channel realization (or, more precisely, the channel's average mutual information) then it does not know at what rate to transmit such that the data will be received correctly. In this case the appropriate capacity definition is capacity with outage. In capacity with outage the transmitter fixes a transmission rate R, and the outage probability associated with R is the probability that the transmitted data will not be received correctly or, equivalently, the probability that the channel $\mathbf{H}$ has mutual information less than R. This probability is given by

$$P_{\text{out}} = p\left(\mathbf{H} : B\log_2 \det\left[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H\right] < R\right). \tag{10.20}$$

This probability is determined by the distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^H$: these eigenvalues are the squares of the singular values of $\mathbf{H}$. The distribution of singular values for matrices is a well-studied problem and is known for common matrices associated with the MIMO channel [17, Sec. 2.1].

As the number of transmit and receive antennas grows large, random matrix theory provides a central limit theorem for the distribution of the singular values of $\mathbf{H}$ [20], resulting in a constant mutual information for all channel realizations. These results were applied to obtain MIMO channel capacity with uncorrelated fading in [21, 22, 23, 24] and with correlated fading in [25, 26, 27]. As an example of this limiting distribution, note that, for fixed M_r, under the ZMSW model the law of large numbers implies that

$$\lim_{M_t\to\infty} \frac{1}{M_t}\mathbf{H}\mathbf{H}^H = \mathbf{I}_{M_r}. \tag{10.21}$$

Substituting this into (10.18) yields that

$$\lim_{M_t\to\infty} I(\mathbf{x};\mathbf{y}) = \lim_{M_t\to\infty} B\log_2 \det\left[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H\right] = B\log_2 \det\left[\mathbf{I}_{M_r} + \rho\mathbf{I}_{M_r}\right] = M_r B\log_2(1+\rho). \tag{10.22}$$

Defining $M = \min(M_t, M_r)$, this implies that as M grows large, the MIMO channel capacity in the absence of CSIT approaches $C = MB\log_2(1+\rho)$ and hence grows linearly in M. Moreover, this linear growth of capacity with M in the asymptotic limit of large M is observed even for a small number of antennas [27].

A MIMO system with a large number of transmit and receive antennas has been coined a *Massive MIMO* system [18, 19]. The benefits of massive MIMO go well beyond linear growth in capacity. That is because, theoretically, in the limit of an asymptotically large number of antennas at the transmitter and receiver, many impairments in wireless communications are removed in a massive MIMO system. In particular, the asymptotically narrow beamwidths at the transmit and receive arrays cause signal reflections to be significantly attenuated. Hence, flat and frequency selective fading disappear because there is no multipath. Path loss also disappears because all of the transmit and receive signal energy is focused within an asymptotically narrow beamwidth. Interference also becomes negligible because most interfering signals arrive outside the asymptotically small angle of the receiver beam. These benefits come at a cost, as Massive MIMO systems are more expensive, larger, and consume more energy than smaller MIMO or SISO systems.

Many MIMO systems preclude having a large number of transmit and receive antennas due to complexity or space constraints. In such systems, we can look at capacity as SNR grows large for a fixed number of transmit and receive antennas. It is shown in [3] that for asymptotically large SNR, capacity also grows linearly with $M = \min(M_t, M_r)$ for any M_t and M_r. Since the ZMSW MIMO channel has $R_{\mathbf{H}} = M = \min(M_t, M_r)$ we see that, for this channel, capacity in the absence of CSIT at high SNRs or with a large number of antennas increases linearly with the number of degrees of freedom in the channel. These results are the main reason for the widespread appeal of MIMO techniques; even if the channel realization is not known at the transmitter, the capacity of ZMSW MIMO channels still grows linearly with the minimum number of transmit and receiver antennas – as long as the channel can be accurately estimated at the receiver. Thus, MIMO channels can provide high data rates without requiring increased signal power or bandwidth.

It is important to note, however, that transmit antennas are not beneficial at very low SNRs; in this regime capacity scales only with the number of receive antennas, independent of the number of transmit antennas. The reason is that, at these low SNRs, the MIMO system is just trying to collect energy rather than trying to exploit all available degrees of freedom. Hence, for the ZMSW MIMO channel, energy can be spread over all transmit antennas or concentrated in just one (or a few) of these antennas in order to achieve capacity [5]. As SNR increases, power ceases to be a limiting factor and capacity becomes limited by the degrees of freedom in the channel.

Although lack of CSIT does not affect the growth rate of capacity relative to M, at least for a large number of antennas, it does complicate demodulation. Specifically, without CSIT the transmission scheme cannot convert the MIMO channel into noninterfering SISO channels. Recall that decoding complexity is exponential in the number of independent symbols transmitted over the multiple transmit antennas, and this number equals the rank of the input covariance matrix.

Our analysis here under perfect CSIR and no CSIT assumes that the channel gain matrix has a ZMSW distribution; that is, it has mean zero and covariance matrix equal to the identity matrix. When the channel has nonzero mean or a nonidentity covariance matrix, there is a spatial bias in the channel that should be exploited by the optimal transmission strategy, so equal power allocation across antennas is no longer optimal [28, 29, 30]. Results in [28, 31] indicate that, when the channel has a dominant mean or covariance direction, *beamforming* (described in Section 10.4 can be used to achieve channel capacity. This is a fortuitous situation, given the simplicity of beamforming.

10.3.2 Fading Channels

Suppose now that the channel gain matrix experiences flat fading, so the gains h_{ij} vary with time. As in the case of the static channel, the capacity depends on what is known about the channel matrix at the transmitter and receiver. With perfect CSIR and CSIT the transmitter can adapt to the channel fading, and its capacity equals the average

over all channel matrix realizations with optimal power allocation. With CSIR and no CSIT, ergodic capacity and capacity with outage are used to characterize channel capacity. These different characterizations are described in more detail in the following sections.

CHANNEL KNOWN AT TRANSMITTER: WATER-FILLING

With CSIT and CSIR, the transmitter optimizes its transmission strategy for each fading channel realization, as in the case of a static channel. The capacity is then just the average of capacities associated with each channel realization, given by (10.11), with power optimally allocated. This average capacity is called the *ergodic* capacity of the channel. There are two possibilities for allocating power under ergodic capacity. A short-term power constraint assumes that the power associated with each channel realization must equal the average power constraint $\bar{P}$. In this case the ergodic capacity becomes

$$\begin{aligned} C &= \mathbf{E}_{\mathbf{H}} \left[\max_{\mathbf{R}_{\mathbf{x}}:\mathrm{Tr}(\mathbf{R}_{\mathbf{x}})=\rho} B \log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R}_{\mathbf{x}}\mathbf{H}^H] \right] \\ &= \mathbf{E}_{\mathbf{H}} \left[\max_{P_i:\Sigma_i P_i \le \bar{P}} \sum_i B \log_2 \left(1 + \frac{P_i \gamma_i}{\bar{P}} \right) \right], \end{aligned} \tag{10.23}$$

where $\gamma_i = \sigma_i^2 \bar{P}/\sigma^2$. A less restrictive constraint is a long-term power constraint, where we can use different powers $P_{\mathbf{H}}$ for different channel realizations $\mathbf{H}$ subject to the average power constraint over all channel realizations $\mathbf{E}_{\mathbf{H}}[P_{\mathbf{H}}] \le \bar{P}$. The ergodic capacity under this assumption is given by

$$\begin{aligned} C &= \max_{\rho_{\mathbf{H}}:\mathbf{E}_{\mathbf{H}}[\rho_{\mathbf{H}}]=\rho} \mathbf{E}_{\mathbf{H}} \left[\max_{\mathbf{R}_{\mathbf{x}}:\mathrm{Tr}(\mathbf{R}_{\mathbf{x}})=\rho_{\mathbf{H}}} B \log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R}_{\mathbf{x}}\mathbf{H}^H] \right] \\ &= \max_{P_{\mathbf{H}}:\mathbf{E}_{\mathbf{H}}[P_{\mathbf{H}}]\le \bar{P}} \mathbf{E}_{\mathbf{H}} \left[\max_{P_i:\Sigma_i P_i \le P_{\mathbf{H}}} \sum_i B \log_2 \left(1 + \frac{P_i \gamma_i}{P_{\mathbf{H}}} \right) \right] \end{aligned} \tag{10.24}$$

for $\gamma_i = \sigma_i^2 P_{\mathbf{H}}/\sigma^2$. The short-term power constraint gives rise to a water-filling in space across the antennas, whereas the long-term power constraint allows for a two-dimensional water-filling across both space and time; this is similar to the frequency–time water-filling associated with the capacity of a time-varying frequency-selective fading channel. The bracketed terms in the second lines of (10.23) and (10.24) are functions of the singular values $\{\sigma_i\}$ of the matrix $\mathbf{H}$. Thus the expectation with respect to $\mathbf{H}$ in these expressions will be based on the distribution of these singular values. This distribution for $\mathbf{H}$ a ZMSW matrix, as well as for other types of matrices, can be found in [17, Sec. 2.1].

CHANNEL UNKNOWN AT TRANSMITTER: ERGODIC CAPACITY AND CAPACITY WITH OUTAGE

Consider now a time-varying channel with random matrix $\mathbf{H}$ that is known at the receiver but not the transmitter. The transmitter assumes a ZMSW distribution for $\mathbf{H}$. The two relevant capacity definitions in this case are ergodic capacity and capacity with outage. Ergodic capacity defines the maximum rate, averaged over all channel realizations, that can be transmitted over the channel for a transmission strategy based only on the distribution of $\mathbf{H}$. This leads to the transmitter optimization problem – that is, finding the optimum input covariance matrix to maximize ergodic capacity subject to the transmit power constraint. Mathematically, the problem is one of characterizing the optimum $\mathbf{R}_{\mathbf{x}}$ to maximize

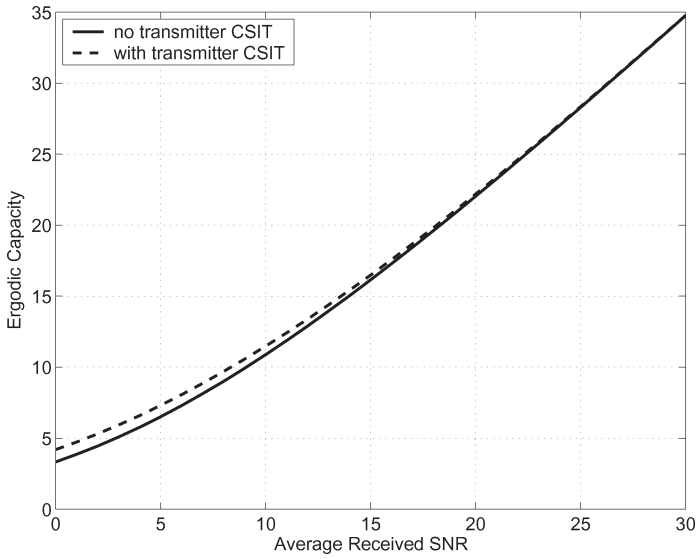


Figure 10.4: Ergodic capacity of 4×4 MIMO channel.

$$C = \max_{\mathbf{R_x}:\mathrm{Tr}(\mathbf{R_x})=\rho} \mathbf{E_H}[B \log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H]]; \tag{10.25}$$

here the expectation is with respect to the distribution on the channel matrix $\mathbf{H}$, which for the ZMSW model is i.i.d. zero-mean circularly symmetric and of unit variance.

As in the case of static channels, the optimum input covariance matrix that maximizes ergodic capacity for the ZMSW model is the scaled identity matrix $\mathbf{R_x} = (\rho/M_t)\mathbf{I}_{M_t}$; that is, the transmit power is divided equally among all the transmit antennas, and independent symbols are sent over the different antennas. Thus the ergodic capacity is given by

$$C = \mathbf{E_H}\left[B \log_2 \det\left[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H\right]\right], \tag{10.26}$$

where, as when CSIT is available, the expectation with respect to $\mathbf{H}$ is based on the distribution of its singular values. The capacity of the static channel grows as $M = \min(M_t, M_r)$ for M large, so this will also be true of the ergodic capacity because it simply averages the static channel capacity. Expressions for the growth rate constant can be found in [5, 32]. When the channel is not ZMSW, capacity depends on the distribution of the singular values for the random channel matrix; these distributions and the resulting ergodic capacity in this more general setting are studied in [16].

The ergodic capacity of a 4×4 MIMO system with i.i.d. complex Gaussian channel gains is shown in Figure 10.4. This figure shows capacity with both transmitter and receiver CSI and with receiver CSI only. There is little difference between the two plots and the difference decreases with SNR, which is also the case for an SISO channel. Comparing the capacity of this channel to that of the SISO fading channel shown in Figure 4.7, we see that the MIMO ergodic capacity is four times larger than the SISO ergodic capacity, which is just as expected since $\min(M_t, M_r) = 4$.

If the channel gain matrix is unknown at the transmitter and the entries are complex Gaussian but not i.i.d., then the channel mean or covariance matrix can be used at the transmitter to increase capacity. The basic idea is to allocate power according to the mean or covariance. This channel model is sometimes referred to as mean or covariance feedback. This model assumes perfect receiver CSI, and the impact of correlated fading depends

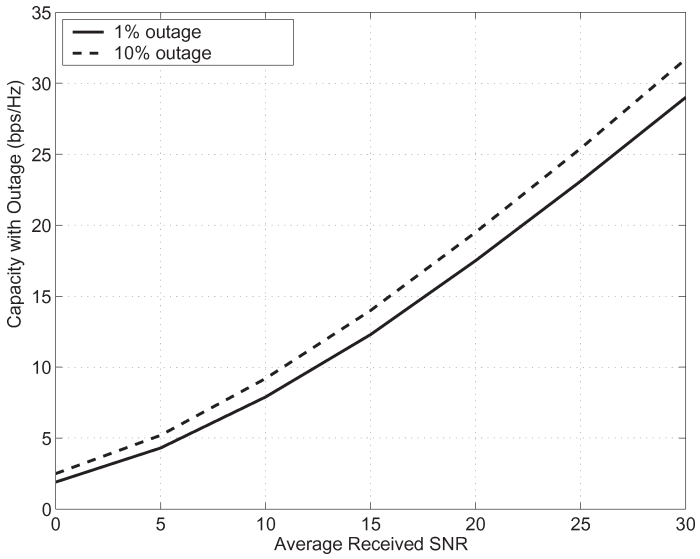


Figure 10.5: Capacity with outage of a 4×4 MIMO channel.

on what is known at the transmitter: if the transmitter knows the channel realization or if it knows *neither* the channel realization nor the correlation structure, then antenna correlation decreases capacity relative to i.i.d. fading. However, if the transmitter knows the correlation structure then, for $M_r = 1$, capacity is increased relative to i.i.d. fading. Details on capacity under these different conditions can be found in [28, 31, 33].

Capacity with outage is defined similarly to the definition for static channels described in Section 10.3.1, although now capacity with outage applies to a slowly varying channel, where the channel matrix $\mathbf{H}$ is constant over a relatively long transmission time and then changes to a new value. As in the static channel case, the channel realization and corresponding channel capacity are not known at the transmitter, yet the transmitter must still fix a transmission rate in order to send data over the channel. For any choice of this rate R, there will be an outage probability associated with R that equals the probability that the transmitted data will not be received correctly. The outage probability is the same as in the static case, given by (10.20). The capacity with outage can sometimes be improved by not allocating power to one or more of the transmit antennas, especially when the outage probability is high [5]. This is because capacity with outage depends on the tail of the probability distribution. With fewer antennas, less averaging takes place and the spread of the tail increases.

The capacity with outage of a 4×4 MIMO system with i.i.d. complex Gaussian channel gains is shown in Figure 10.5 for outage of 1% and 10%. We see that the difference in capacity with outage for these two outage probabilities increases with SNR, which can be explained by the distribution curves for capacity shown in Figure 10.6. These curves show that at low SNRs the distribution is very steep, so that the capacity with outage at 1% is very close to that at 10% outage. At higher SNRs the curves become less steep, leading to more of a capacity difference at different outage probabilities.

NO CSI AT THE TRANSMITTER OR RECEIVER

When there is no CSI at either the transmitter or receiver, the linear growth in capacity as a function of the number of transmit and receive antennas disappears, and in some cases adding additional antennas provides negligible capacity gain. Moreover, channel capacity becomes heavily dependent on the underlying channel model, which makes it difficult to obtain generalizations about capacity growth. For an i.i.d. ZMSW block fading channel it is shown in [34] that increasing the number of transmit antennas by more than the duration of the block does not

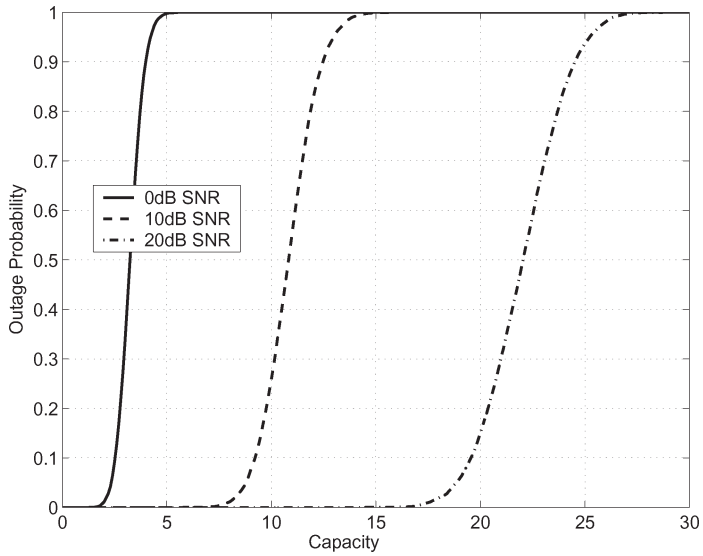


Figure 10.6: Outage probability distribution of a 4×4 MIMO channel.

increase capacity. Thus, there is no data rate increase beyond a certain number of transmit antennas. However, when fading is correlated, additional transmit antennas do increase capacity [35]. The results of [34] were extended in [36] to explicitly characterize capacity and the capacity-achieving transmission strategy for the i.i.d. block fading model in the high-SNR regime. It was shown in [36] that, for i.i.d. block fading channels, capacity grows linearly with the number of channel degrees of freedom.

The capacity of more general fading channel models without transmitter or receiver CSI was investigated in [37, 38, 39]. These works indicate that there are three distinct regions associated with the capacity of slowly varying flat fading channels with no CSIT or CSIR. At low SNRs capacity is limited by noise and grows linearly with the number of channel degrees of freedom. At moderate to high SNRs capacity is limited by estimation error, and its growth is also linear in the number of channel degrees of freedom. At very high SNRs the degrees of freedom cease to matter and capacity grows double logarithmically with SNR. In other words, at very high SNRs there is no multiplexing gain associated with multiple antennas for slowly varying channels without transmitter or receiver CSI.

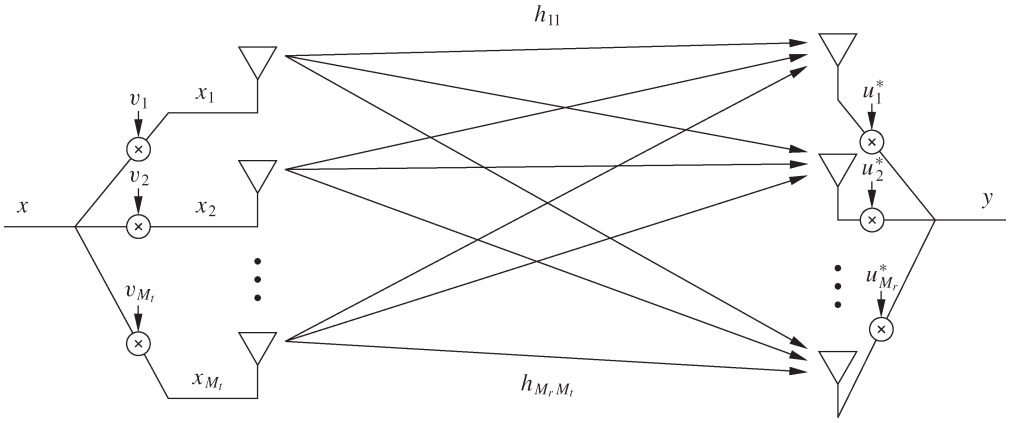


Figure 10.7: MIMO channel with beamforming.

10.4 MIMO Diversity Gain: Beamforming

The multiple antennas at the transmitter and receiver can be used to obtain array and diversity gain (defined in Section 10.1) instead of capacity gain. In this setting the same symbol – weighted by a complex scale factor – is sent over each transmit antenna, so that the input covariance matrix has unit rank. This scheme is also referred to as *MIMO beamforming*. A beamforming strategy corresponds to the precoding and combining matrices (described in Section 10.2) being just column vectors: $\mathbf{V} = \mathbf{v}$ and $\mathbf{U} = \mathbf{u}$, as shown in Figure 10.7. As indicated in the figure, the transmit symbol x is sent over the ith antenna with weight v_i. On the receive side, the signal received on the ith antenna is weighted by u_i^*. Both transmit and receive weight vectors are normalized so that $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. The resulting received signal is given by

$$y = \mathbf{u}^H \mathbf{H} \mathbf{v} x + \mathbf{u}^H \mathbf{n}, \tag{10.27}$$

where if $\mathbf{n} = (n_1, \ldots, n_{M_r})$ has i.i.d. elements then the statistics of $\mathbf{u}^H \mathbf{n}$ are the same as the statistics for each of these elements.

Beamforming provides diversity and array gain via coherent combining of the multiple signal paths. Channel knowledge at the receiver is assumed, since this is required for coherent combining. The performance gain then depends on whether or not the channel is known at the transmitter. When the channel matrix $\mathbf{H}$ is known, the received SNR is optimized by choosing $\mathbf{u}$ and $\mathbf{v}$ as the principal left and right singular vectors of the channel matrix $\mathbf{H}$. That is, for $\sigma_1 = \sigma_{\max}$ the maximum singular value of $\mathbf{H}$, $\mathbf{u}$ and $\mathbf{v}$ are (respectively) the first columns of $\mathbf{U}$ and $\mathbf{V}$. The corresponding received SNR can be shown to equal $\gamma = \sigma_{\max}^2 \rho$, where $\sigma_{\max}$ is the largest singular value of $\mathbf{H}$ [9, 40]. The resulting capacity is $C = B \log_2(1 + \sigma_{\max}^2 \rho)$, corresponding to the capacity of a SISO channel with channel power gain $\sigma_{\max}^2$. For $\mathbf{H}$ a ZMSW matrix, it can be shown [9, Chap. 5.4.4] that the array gain of beamforming diversity is between $\max(M_t, M_r)$ and $M_t M_r$ and that the diversity gain is $M_t M_r$.

When the channel is not known to the transmitter, for $M_t = 2$ the Alamouti scheme described in Section 7.3.2 can be used to extract an array gain of M_r and the maximum diversity gain of $2M_r$ [9, Chap. 5.4.3]. For $M_t > 2$, full diversity gain can also be obtained using space-time block codes, as described in Section 10.7.3. Although beamforming has a reduced capacity relative to the capacity-achieving transmit precoding and receiver combining matrices, the demodulation complexity with beamforming is on the order of $|\mathcal{X}|$ instead of $|\mathcal{X}|^{R_{\mathrm{H}}}$. An even simpler strategy is to use MRC at either the transmitter or receiver and antenna selection on the other end: this was analyzed in [41].

Example 10.3: Consider a MIMO channel with gain matrix

$$\mathbf{H} = \begin{bmatrix} .7 & .9 & .8 \\ .3 & .8 & .2 \\ .1 & .3 & .9 \end{bmatrix}. \tag{10.28}$$

Find the capacity of this channel under beamforming, given channel knowledge at the transmitter and receiver, $B = 100$ kHz, and $\rho = 10$ dB.

Solution: The largest singular value of $\mathbf{H}$ is $\sigma_{\max} = \sqrt{\lambda_{\max}}$, where $\lambda_{\max}$ is the maximum eigenvalue of

$$\mathbf{H}\mathbf{H}^H = \begin{bmatrix} 1.94 & 1.09 & 1.06 \\ 1.09 & .77 & .45 \\ 1.06 & .45 & .91 \end{bmatrix}. \tag{10.29}$$

The largest eigenvalue of this matrix is $\lambda_{\max} = 3.17$. Thus, $C = B\log_2(1 + \lambda_{\max}\rho) = 10^5 \log_2(1 + 31.7) = 503$ kbps.

10.5 Diversity–Multiplexing Trade-offs

The previous sections suggest two mechanisms for utilizing multiple antennas to improve wireless system performance. One option is to obtain capacity gain by decomposing the MIMO channel into parallel channels and multiplexing different data streams onto these channels. This capacity gain is also referred to as a *multiplexing gain*. However, the SNR associated with each of these channels depends on the singular values of the channel matrix. In capacity analysis this is taken into account by assigning a relatively low rate to these channels. However, practical signaling strategies for these channels will typically have poor performance unless powerful channel coding techniques are employed. Alternatively, beamforming can be used, where the channel gains are coherently combined to obtain a very robust channel with high diversity gain. It is not necessary to use the antennas purely for multiplexing or diversity. Some of the space-time dimensions can be used for diversity gain and the remaining dimensions used for multiplexing gain. This gives rise to a fundamental design question in MIMO systems: Should the antennas be used for diversity gain, multiplexing gain, or both?

The diversity–multiplexing trade-off or, more generally, the trade-off between data rate, probability of error, and complexity for MIMO systems has been extensively studied in the literature – both from a theoretical perspective and in terms of practical space-time code designs [42, 43, 44, 45]. This work has primarily focused on block fading channels with receiver CSI only, since when both transmitter and receiver know the channel the trade-off is relatively straightforward: antenna subsets can first be grouped for diversity gain, and then the multiplexing gain corresponds to the new channel with reduced dimension due to the grouping. For the block fading model with receiver CSI only, as the blocklength grows asymptotically large, full diversity gain and full multiplexing gain (in terms of capacity with outage) can be obtained simultaneously with reasonable complexity by encoding diagonally across antennas [3, 46, 47]. An example of this type of encoding is D-BLAST, described in Section 10.7.4. For finite blocklengths it is not possible to achieve full diversity and full multiplexing gain simultaneously, in which case there is a trade-off between these gains. A simple characterization of this trade-off is given in [45] for block fading channels in the limit of asymptotically high SNR. In this analysis a transmission scheme is said to achieve multiplexing gain r and diversity gain d if the data rate (bps) per unit hertz, $R(\mathrm{SNR})$, and probability of error, $P_e(\mathrm{SNR})$, as functions of SNR satisfy

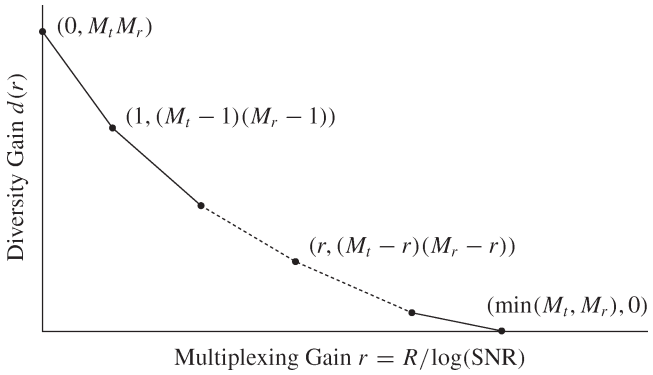


Figure 10.8: Diversity–multiplexing trade-off for high-SNR block fading.

$$\lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\log_2 \mathrm{SNR}} = r \tag{10.30}$$

and

$$\lim_{\mathrm{SNR}\to\infty} \frac{\log P_e(\mathrm{SNR})}{\log \mathrm{SNR}} = -d, \tag{10.31}$$

where the log in (10.31) can be in any base.[5] For each r, the optimal diversity gain $d_{\mathrm{opt}}(r)$ is the maximum diversity gain that can be achieved by any scheme. It is shown in [45] that if the fading blocklength $T \geq M_t + M_r - 1$ then

$$d_{\mathrm{opt}}(r) = (M_t - r)(M_r - r), \quad 0 \leq r \leq \min(M_t, M_r). \tag{10.32}$$

The function (10.32) is plotted in Figure 10.8. This figure implies that (i) if we use all transmit *and* receive antennas for diversity then we get full diversity gain $M_t M_r$ and (ii) we can use some of these antennas to increase data rate at the expense of diversity gain.

It is also possible to adapt the diversity and multiplexing gains relative to channel conditions. Specifically, in poor channel states more antennas can be used for diversity gain, whereas in good states more antennas can be used for multiplexing. Adaptive techniques that change antenna use to trade off diversity and multiplexing based on channel conditions have been investigated in [48, 49, 50].

Example 10.4: Let the multiplexing and diversity parameters r and d be as defined in (10.30) and (10.31). Suppose that r and d approximately satisfy the diversity–multiplexing trade-off $d_{\mathrm{opt}}(r) = (M_t - r)(M_r - r)$ at any large finite SNR. For an $M_t = M_r = 8$ MIMO system with an SNR of 15 dB, if we require a data rate per unit hertz of $R = 15$ bps then what is the maximum diversity gain the system can provide?
Solution: With SNR $= 15$ dB, to get $R = 15$ bps we require $r \log_2(10^{1.5}) = 15$, which implies that $r = 3.01$. Therefore, three of the antennas are used for multiplexing and the remaining five are used for diversity. The maximum diversity gain is then $d_{\mathrm{opt}}(r) = (M_t - r)(M_r - r) = (8 - 3)(8 - 3) = 25$.

10.6 MIMO Detection Algorithms

When the transmitter and receiver have perfect CSI, transmit precoding and receiver combining can be used to transform the MIMO channel into separate SISO channels, for which the optimal symbol detection techniques of Chapter 5.1.3 can be used. In the absence of CSIT and the associated transmit precoder, the MIMO receiver must detect the vector of transmitted symbols associated with the received signal. In this section we describe algorithms used at a MIMO receiver for this vector detection. In particular, we consider a MIMO receiver that receives vector $\mathbf{y}$ given by (10.2):

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{10.33}$$

where the input vector $\mathbf{x} \in \mathcal{X}^{M_t}$ and, for $\mathbf{C}$ the set of all complex numbers, the complex received vector $\mathbf{y} \in \mathbf{C}^{M_r}$ and $\mathbf{H} \in \mathbf{C}^{M_r \times M_t}$. We assume a normalized power ρ such that $\mathrm{E}[\mathbf{x}\mathbf{x}^H] = \frac{\rho}{M_t}\mathbf{I}$ such that the complex noise n_i has real and imaginary part distributed as $\mathcal{N}(0, 1)$. The receiver estimates $\hat{\mathbf{x}}(\mathbf{y})$, an estimate for the transmitted vector $\mathbf{x}$, based on its knowledge of the channel matrix $\mathbf{H}$, $\mathcal{X}$, and the received vector $\mathbf{y}$.

[5] The base of the log cancels out of the expression because (10.31) is the ratio of two logs with the same base.

We consider the following detection algorithms: maximum-likelihood (ML), the linear algorithms of zero-forcing (ZF) and linear minimum mean-square error (L-MMSE), and an approximation to ML called sphere decoding (SD), where the radius of the sphere provides a tradeoff between performance and complexity. Sphere decoding becomes ML for an asymptotically large radius, so reducing the radius reduces both complexity and performance relative to ML.

10.6.1 Maximum Likelihood (ML) Detection

As in scalar signal detection, the ML MIMO detection algorithm minimizes the probability of symbol detection error (assuming equiprobable $\mathbf{x}$). In particular, the maximum likelihood detector with i.i.d. Gaussian noise at the receiver antennas solves the following problem:

$$\hat{\mathbf{x}}(\mathbf{y}) = \text{argmin}_{\mathbf{x}\in\mathcal{X}^{M_t}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \tag{10.34}$$

The minimization is over $\mathbf{x} \in \mathcal{X}^{M_t}$, i.e. over all possible transmitted vectors. Unfortunately, solving this problem involves computing the objective function for all $\mathcal{X}^{M_t}$ potential values of $\mathbf{x}$. Hence the ML detector has prohibitive (exponential in M_t) complexity.

10.6.2 Linear Receivers: ZF and MMSE

The complexity of linear detectors is the same as the complexity of inverting or factorizing a matrix of dimensions $M_r \times M_t$, hence the name. They work by spatially decoupling the effects of the channel by a process known as MIMO equalization. This involves multiplying $\mathbf{y}$ with a MIMO equalization matrix $\mathbf{A} \in \mathbf{C}^{M_t \times M_r}$ to get $\tilde{\mathbf{x}}(\mathbf{y}) \in \mathbf{C}^{M_t}$. To get the estimate $\hat{\mathbf{x}}(\mathbf{y})$ from $\tilde{\mathbf{x}}(\mathbf{y})$, we perform coordinatewise decoding of $\tilde{\mathbf{x}}(\mathbf{y})$. The coordinatewise decoding operation is given by:

$$\hat{\mathbf{x}}(\mathbf{y})_i = \text{argmin}_{s\in\mathcal{X}} |\tilde{\mathbf{x}}(\mathbf{y})_i - s| \text{ for all } i, \tag{10.35}$$

i.e., it maps each coordinate to the closest constellation point.

A common way to obtain $\tilde{\mathbf{x}}(\mathbf{y})$ from $\mathbf{y}$ is through zero-forcing (ZF), which solves the following problem:

$$\tilde{\mathbf{x}}(\mathbf{y}) = \text{argmin}_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \tag{10.36}$$

Comparing with (10.34), we note that the constellation constraints on $\mathbf{x}$ have been removed. This significantly reduces the complexity. The solution is given by

$$\tilde{\mathbf{x}}(\mathbf{y}) = \mathbf{H}^\dagger \mathbf{y}, \tag{10.37}$$

where $\mathbf{H}^\dagger$ is the **pseudo inverse** of $\mathbf{H}$. Note that $\mathbf{H}^\dagger$ is just $\mathbf{H}^{-1}$ if the matrix is square and invertible. When $M_t \leq M_r$, and there are M_t linearly independent columns in $\mathbf{H}$, which is often the case in typical propagation environments, the pseudo inverse (sometimes called the Moore-Penrose pseudoinverse) is given by

$$\mathbf{H}^\dagger = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H. \tag{10.38}$$

When $M_t \geq M_r$ and there are M_r linearly independent rows in $\mathbf{H}$, we have

$$\mathbf{H}^\dagger = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}. \tag{10.39}$$

In general, $\mathbf{H}^\dagger$ can be derived by singular value decomposition (SVD). Let $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ be the SVD of $\mathbf{H}$, then we have

$$\mathbf{H}^\dagger = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^H. \tag{10.40}$$

The complexity of obtaining $\mathbf{H}^\dagger$ from $\mathbf{H}$ is roughly cubic in M_t for a square matrix. However obtaining $\hat{\mathbf{x}}(\mathbf{y})$ from $\tilde{\mathbf{x}}(\mathbf{y})$ is done in a time linear in M_t.

Zero forcing can cause noise amplification if the minimum singular value of $\mathbf{H}$ is much smaller than its maximum value. This may be quantified by the notion of the condition number of the matrix $\mathbf{H}$. The condition number of the matrix $\mathbf{H}$ is a measure of the relative magnitudes of the singular values of $\mathbf{H}$. It is defined as the ratio between the largest and the smallest singular values of $\mathbf{H}$. When the condition number is unity or close to unity, the matrix is said to be well conditioned. When the condition number is large, the matrix is ill conditioned. When the matrix is ill-conditioned, noise is significantly amplified in the spatial channels associated with small singular values. This problem can be compensated for using the minimum mean squared error (MMSE) detector, which adds a regularization term to the objective function in (10.36).

$$\tilde{\mathbf{x}}(\mathbf{y}) = \text{argmin}_{\mathbf{x}}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda\|\mathbf{x}\|^2, \tag{10.41}$$

for some $\lambda > 0$. The solution to this is given by

$$\tilde{\mathbf{x}}(\mathbf{y}) = (\mathbf{H}^H\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^H\mathbf{y} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \lambda\mathbf{I})^{-1}\mathbf{y}. \tag{10.42}$$

For $\lambda = \frac{M_t}{\rho}$, this is called the **L-MMSE** detector since it minimizes the mean squared error in the estimate of $\mathbf{x}$ among all linear detectors, i.e., it solves the following problem:

$$\begin{aligned}\tilde{\mathbf{x}}(\mathbf{y}) = \text{argmin}_{\mathbf{s}}\ \mathrm{E}_{\mathbf{x},\mathbf{n}}[\|\mathbf{s} - \mathbf{x}\|^2], \\ \text{such that } \mathbf{s} = \mathbf{A}\mathbf{y} + \mathbf{b},\end{aligned} \tag{10.43}$$

for any $M_t \times M_r$ matrix $\mathbf{A}$ and vector $\mathbf{b} \in \mathbf{C}^{M_t}$. Note that the minimization is only over all affine functions of $\mathbf{y}$, which is parameterized by $\mathbf{A}$ and $\mathbf{b}$. The expectation is over the randomness in $\mathbf{x}$ and $\mathbf{n}$ (the channel matrix $\mathbf{H}$ is assumed to be known and non random). If $\mathbf{x}$ were to be Gaussian (instead of being from discrete constellation points), this is also the MMSE detector.

Compared to the ML detector, both the ZF and MMSE linear detectors are simpler to implement, but their BER performances are worse. A nonlinear technique that allows for better performance than a linear detector and a simple mechanism to trade off between complexity and performance is the sphere decoder, which we describe in the next subsection.

10.6.3 Sphere Decoders (SD)

The sphere decoder trades off performance versus complexity by controlling a parameter r. By choosing a large enough r, the performance of SD approaches that of the ML detector. For small r, the search space (and hence complexity) of the SD is much smaller than that of the ML detector, but it suffers a performance degradation as a result. The sphere detector exploits the following factorization of the matrix $\mathbf{H}$:

$$\mathbf{H} = \mathbf{Q}\mathbf{R}, \tag{10.44}$$

where $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper triangular. Since the squared-distance norm does not change under multiplication by a unitary matrix $\mathbf{Q}^H$, we have

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 = \|\mathbf{Q}^H\mathbf{y} - \mathbf{Q}^H\mathbf{Q}\mathbf{R}\mathbf{x}\|^2 = \|\tilde{\mathbf{y}} - \mathbf{R}\mathbf{x}\|^2 = \sum_{i=1}^{M_t}|\tilde{\mathbf{y}}_i - \sum_{j=i}^{M_t}\mathbf{R}_{i,j}\mathbf{x}_j|^2, \tag{10.45}$$

where $\tilde{\mathbf{y}} \triangleq \mathbf{Q}^H\mathbf{y}$, and we use the upper triangular nature of $\mathbf{R}$ in the last step. The sphere decoder solves the following problem:

$$\hat{\mathbf{x}}(\mathbf{y}) = \text{argmin}_{\mathbf{x}\in\mathcal{X}^{M_t}, \|\mathbf{Q}^H\mathbf{y}-\mathbf{R}\mathbf{x}\|\leq r}\|\mathbf{Q}^H\mathbf{y} - \mathbf{R}\mathbf{x}\|^2. \tag{10.46}$$

Note that choosing $r = \infty$ gives us the ML decoder. For a smaller r, the solver can exploit the upper triangular nature of $\mathbf{R}$ to "prune" many candidate solutions (using depth-first-search or breadth-first-search or a combination of the two), thereby reducing the detection complexity significantly. One surprising property of the SD is that if it finds a valid solution, it is the same solution that the ML detector would have returned.

10.7 MIMO Space-Time Codes

In this section we consider MIMO systems designed for quasi-static channels, where the channel is unknown at the transmitter and assumed constant over a block of T symbol times. The optimal transmission scheme for such channels is designed over both space (via the multiple antennas) and time (via T uses of the channel). Such transmission schemes are referred to as MIMO *space-time codes*.

In MIMO space-time the channel input and output are matrices with dimensions corresponding to space (antennas) and time (symbol blocks). Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T]$ denote the $M_t \times T$ channel input matrix with ith column $\mathbf{x}_i$ equal to the vector channel input over the ith transmission time. Let $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T]$ denote the $M_r \times T$ channel output matrix with ith column $\mathbf{y}_i$ equal to the vector channel output over the ith transmission time, and let $\mathbf{N} = [\mathbf{n}_1, \ldots, \mathbf{n}_T]$ denote the $M_r \times T$ noise matrix with ith column $\mathbf{n}_i$ equal to the receiver noise vector over the ith transmission time. With this matrix representation, the input–output relationship over all T blocks becomes

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}. \tag{10.47}$$

The design of space-time codes is typically based on optimizing a performance criteria. A common criteria is the pairwise error probability, which is discussed in the next section.

10.7.1 Pairwise Error Probability

Assume a space-time code where the receiver has knowledge of the channel matrix $\mathbf{H}$. Under maximum likelihood detection it can be shown using similar techniques as in the scalar (Chapter 5) or vector (Chapter 8) case that, given received matrix $\mathbf{Y}$, the ML transmit matrix $\hat{\mathbf{X}}$ satisfies

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\mathcal{X}^{M_t\times T}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_F^2 = \arg\min_{\mathbf{X}\in\mathcal{X}^{M_t\times T}} \sum_{i=1}^{T} \|\mathbf{y}_i - \mathbf{H}\mathbf{x}_i\|^2, \tag{10.48}$$

where $\|\mathbf{A}\|_F$ denotes the Frobenius norm[6] of the matrix $\mathbf{A}$ and the minimization is taken over all possible space-time input matrices $\mathcal{X}^{M_t\times T}$. The pairwise error probability for mistaking a transmit matrix $\mathbf{X}$ for another matrix $\hat{\mathbf{X}}$, denoted as $p(\hat{\mathbf{X}} \to \mathbf{X})$, depends only on the distance between the two matrices after transmission through the channel and the noise power σ^2; that is,

$$p(\hat{\mathbf{X}} \to \mathbf{X}) = Q\left(\sqrt{\frac{\|\mathbf{H}(\mathbf{X} - \hat{\mathbf{X}})\|_F^2}{2\sigma^2}}\right). \tag{10.49}$$

[6]The Frobenius norm of a matrix is the square root of the sum of the square of its elements.

We now derive an upper bound on the pairwise error probability following the derivation of [9, Chap. 6.3.2]. Let $\mathbf{D_X} = \mathbf{X} - \hat{\mathbf{X}}$ denote the difference matrix between $\mathbf{X}$ and $\hat{\mathbf{X}}$. Applying the Chernoff bound to (10.49), we have

$$p(\hat{\mathbf{X}} \to \mathbf{X}) \leq \exp\left[-\frac{\|\mathbf{H}\mathbf{D_X}\|_F^2}{4\sigma^2}\right]. \tag{10.50}$$

Let $\mathbf{h}_i$ denote the ith row of $\mathbf{H}, i = 1, \ldots, M_r$. Then

$$\|\mathbf{H}\mathbf{D_X}\|_F^2 = \sum_{i=1}^{M_r} \mathbf{h}_i \mathbf{D_X} \mathbf{D_X}^H \mathbf{h}_i^H. \tag{10.51}$$

Let $\mathcal{H} = \text{vec}(\mathbf{H}^T)^T$, where vec($\mathbf{A}$) is defined as the vector that results from stacking the columns of matrix $\mathbf{A}$ on top of each other to form a vector.[7] Hence $\mathcal{H}^T$ is a vector of length $M_r M_t$. Also define $\mathcal{D_X} = \mathbf{I}_{M_r} \otimes \mathbf{D_X}$, where $\otimes$ denotes the Kronecker product. With these definitions,

$$\|\mathbf{H}\mathbf{D_X}\|_F^2 = \|\mathcal{H}\mathcal{D_X}\|_F^2. \tag{10.52}$$

Substituting (10.52) into (10.50) and taking the expectation relative to all possible channel realizations yields

$$p(\mathbf{X} \to \hat{\mathbf{X}}) \leq \left(\det\left[\mathbf{I}_{M_t M_r} + \frac{1}{4\sigma^2}\mathbf{E}[\mathcal{D}_\mathbf{X}^H \mathcal{H}^H \mathcal{H} \mathcal{D_X}]\right]\right)^{-1}. \tag{10.53}$$

Suppose that the channel matrix $\mathbf{H}$ is random and spatially white, so that its entries are i.i.d. zero-mean, unit-variance, complex Gaussian random variables. Then taking the expectation in (10.53) yields

$$p(\mathbf{X} \to \hat{\mathbf{X}}) \leq \left(\frac{1}{\det[\mathbf{I}_{M_t} + .25\rho\mathbf{\Delta}]}\right)^{M_r}, \tag{10.54}$$

where $\mathbf{\Delta} = (1/P)\mathbf{D_X}\mathbf{D_X}^H$. This simplifies to

$$p(\mathbf{X} \to \hat{\mathbf{X}}) \leq \prod_{k=1}^{R_\Delta} \left(\frac{1}{1 + .25\gamma\lambda_k(\mathbf{\Delta})}\right)^{M_r}; \tag{10.55}$$

here $\gamma = P/\sigma^2 = \rho$ is the SNR per input symbol and $\lambda_k(\mathbf{\Delta})$ is the kth nonzero eigenvalue of $\mathbf{\Delta}, k = 1, \ldots, R_\mathbf{\Delta}$, where $R_\mathbf{\Delta}$ is the rank of $\mathbf{\Delta}$. In the high-SNR regime (i.e., for $\gamma \gg 1$) this simplifies to

$$p(\mathbf{X} \to \hat{\mathbf{X}}) \leq \left(\prod_{k=1}^{R_\mathbf{\Delta}} \lambda_k(\Delta)\right)^{-M_r} (.25\gamma)^{-R_\mathbf{\Delta} M_r}. \tag{10.56}$$

This equation gives rise to the main criteria for design of space-time codes, described in the next section.

10.7.2 Rank and Determinant Design Criteria

The pairwise error probability in (10.56) indicates that the probability of error decreases as γ^{-d} for $d = R_\mathbf{\Delta} M_r$. Thus, $R_\mathbf{\Delta} M_r$ is the diversity gain of the space-time code. The maximum diversity gain possible through coherent combining of M_t transmit and M_r receive antennas is $M_t M_r$. Thus, to obtain this maximum diversity gain, the space-time code must be designed such that the $M_t \times M_t$ difference matrix $\mathbf{\Delta}$ between any two codewords has full rank equal to M_t. This design criterion is referred to as the *rank criterion*.

[7] So for the $M \times N$ matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_N]$, where $\mathbf{a}_i$ is a vector of length M, $\text{vec}(\mathbf{A}) = [\mathbf{a}_1^T, \ldots, \mathbf{a}_N^T]^T$ is a vector of length MN.

The coding gain associated with the pairwise error probability in (10.56) depends on the term $(\prod_{k=1}^{R_{\mathbf{\Delta}}} \lambda_k(\mathbf{\Delta}))^{M_r}$. Thus, a high coding gain is achieved by maximizing the minimum of the determinant of $\mathbf{\Delta}$ over all input matrix pairs $\mathbf{X}$ and $\hat{\mathbf{X}}$. This criterion is referred to as the *determinant criterion*.

The rank and determinant criteria were first developed in [43, 51, 52]. These criteria are based on the pairwise error probability associated with different transmit signal matrices – rather than the binary domain of traditional codes–and hence often require computer searches to find good codes [53, 54]. A general binary rank criterion was developed in [55] to provide a better construction method for space-time codes.

10.7.3 Space-Time Trellis and Block Codes

The rank and determinant criteria have been primarily applied to the design of space-time trellis codes (STTCs), which are an extension of conventional trellis codes to MIMO systems [9, 52]. They are described using a trellis and are decoded using ML sequence estimation via the Viterbi algorithm. STTCs can extract excellent diversity and coding gain, but the complexity of decoding increases exponentially with the diversity level and transmission rate [56]. Space-time block codes (STBCs) are an alternative space-time code that can also extract excellent diversity gain with linear receiver complexity. Interest in STBCs was initiated by the Alamouti code described in Section 7.3.2, which obtains full diversity order with linear receiver processing for a two-antenna transmit system. This scheme was generalized in [57] to STBCs that achieve full diversity order with an arbitrary number of transmit antennas. However, while these codes achieve full diversity order, they do not provide coding gain and thus have inferior performance to STTCs, which achieve both full diversity gain as well as coding gain. Added coding gain for STTCs as well as STBCs can be achieved by concatenating these codes either in serial or in parallel with an outer channel code to form a turbo code [58, 59, 60]. The linear complexity of the STBC designs in [57] result from making the codes orthogonal along each dimension of the code matrix. A similar design premise is used in [61] to design *unitary space-time modulation* schemes for block fading channels when neither the transmitter nor the receiver has CSI. More comprehensive treatments of space-time coding can be found in [9, 56, 62, 63] and the references therein.

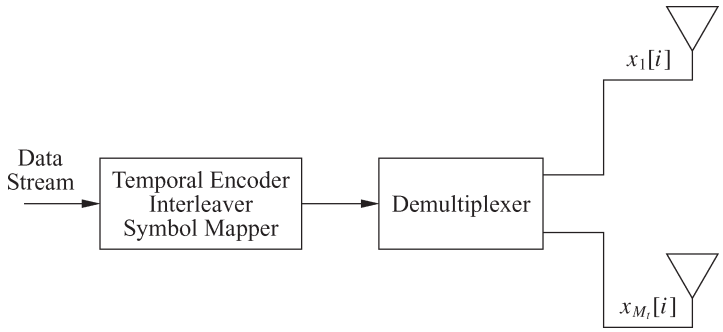


Figure 10.9: Spatial multiplexing with serial encoding.

10.7.4 Spatial Multiplexing and BLAST Architectures

The basic premise of spatial multiplexing is to send M_t independent symbols per symbol period using the dimensions of space and time. To obtain full diversity order, an encoded bit stream must be transmitted over all M_t transmit antennas. This can be done through a serial encoding, illustrated in Figure 10.9. With serial encoding the bit stream is temporally encoded over the channel blocklength T to form the codeword $[x_1, \ldots, x_T]$. The codeword is interleaved and mapped to a constellation point, then demultiplexed onto the different antennas. The first M_t

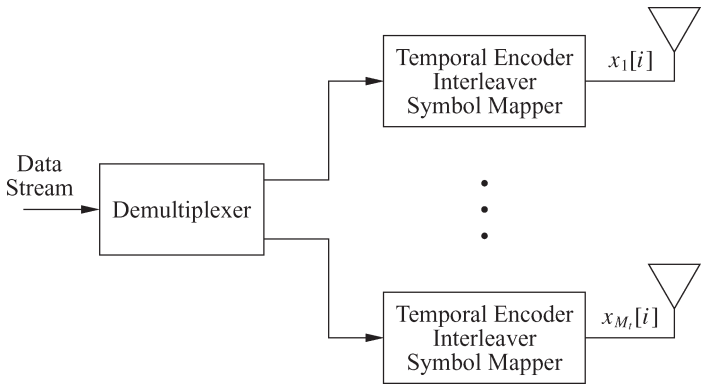


Figure 10.10: Spatial multiplexing with parallel encoding: V-BLAST.

symbols are transmitted from the M_t antennas over the first symbol time, the next M_t symbols are transmitted from the antennas over the next symbol time, and this process continues until the entire codeword has been transmitted. We denote the symbol sent over the kth antenna at time i as $x_k[i]$. If a codeword is sufficiently long, it is transmitted over all M_t transmit antennas and received by all M_r receive antennas, resulting in full diversity gain. However, the codeword length T required to achieve this full diversity is $M_t M_r$, and decoding complexity grows exponentially with this codeword length. This high level of complexity makes serial encoding impractical.

A simpler method to achieve spatial multiplexing, pioneered at Bell Laboratories as one of the Bell Labs Layered Space Time (BLAST) architectures for MIMO channels [3], is parallel encoding, illustrated in Figure 10.10. With parallel encoding the data stream is demultiplexed into M_t independent streams. Each of the resulting substreams is passed through a SISO temporal encoder with blocklength T, interleaved, mapped to a signal constellation point, and transmitted over its corresponding transmit antenna. Specifically, the kth SISO encoder generates the codeword $\{x_k[i],\, i = 1, \ldots, T\}$, which is transmitted sequentially over the kth antenna. This process can be considered to be the encoding of the serial data into a vertical vector and hence is also referred to as *vertical encoding* or V-BLAST [64]. Vertical encoding can achieve at most a diversity order of M_r, since each coded symbol is transmitted from one antenna and received by M_r antennas. This system has a simple encoding complexity that is linear in the number of antennas. However, optimal decoding still requires joint detection of the codewords from each of the transmit antennas, since all transmitted symbols are received by all the receive antennas. It was shown in [65] that the receiver complexity can be significantly reduced through the use of *symbol interference cancellation*, as shown in Figure 10.11. This cancellation, which exploits the synchronicity of the symbols transmitted from each antenna, works as follows. First the M_t transmitted symbols are ordered in terms of their received SNR. An estimate of the received symbol with the highest SNR is made while treating all other symbols as noise. This estimated symbol is subtracted out, and the symbol with the next highest SNR estimated while treating the remaining symbols as noise. This process repeats until all M_t transmitted symbols have been estimated. After cancelling out interfering symbols, the coded substream associated with each transmit antenna can be individually decoded, resulting in a receiver complexity that is linear in the number of transmit antennas. In fact, coding is not even needed with this architecture, and data rates of 20–40 bps/Hz with reasonable error rates were reported in [64] using uncoded V-BLAST.

The simplicity of parallel encoding and the diversity benefits of serial encoding can be obtained by using a creative combination of the two techniques called *diagonal encoding* or D-BLAST [3], illustrated in Figure 10.12. In D-BLAST, the data stream is first parallel encoded. However, rather than transmitting each codeword with one antenna, the codeword symbols are rotated across antennas, so that a codeword is transmitted by all M_t antennas. The operation of the stream rotation is shown in Figure 10.13. Suppose the ith encoder generates the codeword

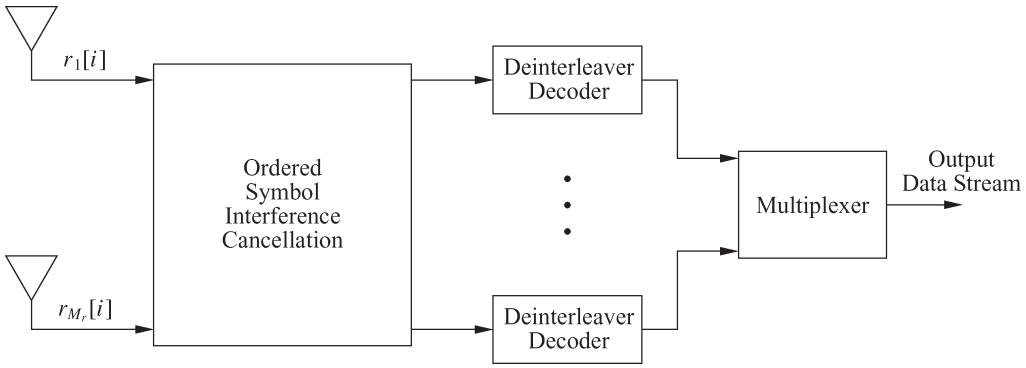


Figure 10.11: V-BLAST receiver with linear complexity.

$\mathbf{x}_i = [x_i[1], \ldots, x_i[T]]$. The stream rotator transmits each symbol on a different antenna, so $x_i[1]$ is sent on antenna 1 over symbol time i, $x_i[2]$ is sent on antenna 2 over symbol time $i+1$, and so forth. If the code blocklength T exceeds M_t then the rotation begins again on antenna 1. As a result, the codeword is spread across all spatial dimensions. Transmission schemes based on D-BLAST can achieve full $M_t M_r$ diversity gain if the temporal coding with stream rotation is capacity-achieving (Gaussian codewords of infinite blocklength T) [9, Chap. 6.3.5]. Moreover, the D-BLAST system can achieve maximum capacity with outage if the wasted space-time dimensions along the diagonals are neglected [9, Chap. 12.4.1]. Receiver complexity is also linear in the number of transmit antennas, since the receiver decodes each diagonal code independently. However, this simplicity comes at a price, as the efficiency loss of the wasted space-time dimensions (illustrated in Figure 10.13) can be large if the frame size is not appropriately chosen.

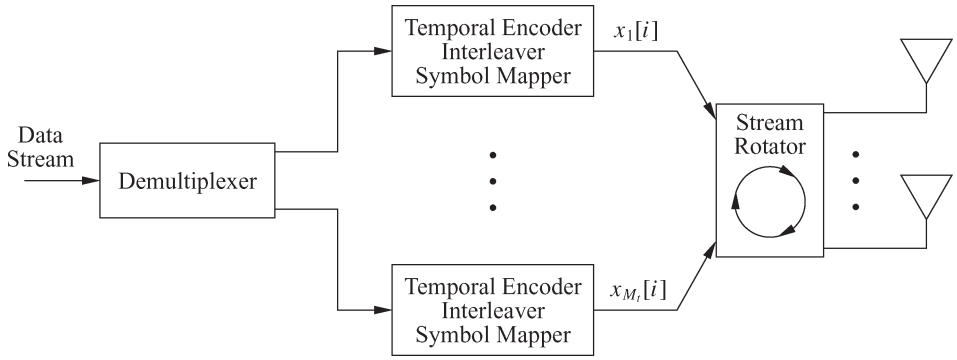


Figure 10.12: Diagonal encoding with stream rotation.

10.8 Frequency-Selective MIMO Channels

When the MIMO channel bandwidth is large relative to the channel's multipath delay spread, the channel suffers from intersymbol interference; this is similar to the case of SISO channels. There are two approaches to dealing with ISI in MIMO channels. A channel equalizer can be used to mitigate the effects of ISI. However, the equalizer is much more complex in MIMO channels because the channel must be equalized over both space and time. Moreover, when the equalizer is used in conjunction with a space-time code, the nonlinear and noncausal nature of the code further complicates the equalizer design. In some cases the structure of the code can be used to convert

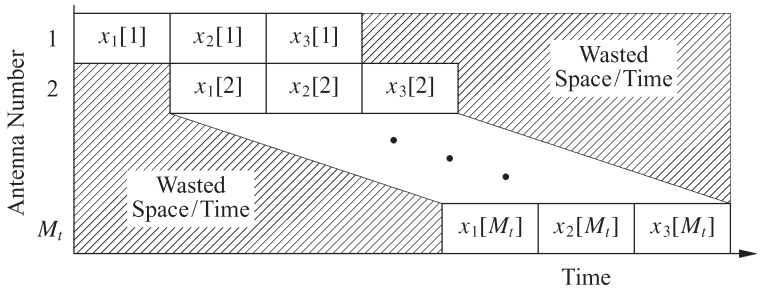


Figure 10.13: Stream rotation.

the MIMO equalization problem to a SISO problem for which well-established SISO equalizer designs can be used [66, 67, 68].

An alternative to equalization in frequency-selective fading is multicarrier modulation or orthogonal frequency division multiplexing (OFDM). OFDM techniques for SISO channels are described in Chapter 12: the main premise is to convert the wideband channel into a set of narrowband subchannels that exhibit only flat fading. Applying OFDM to MIMO channels results in a parallel set of narrowband MIMO channels, and the space-time modulation and coding techniques just described for a single MIMO channel are applied to this parallel set. MIMO frequency-selective fading channels exhibit diversity across space, time, and frequency, so ideally all three dimensions should be fully exploited in the signaling scheme.

Chapter 10 Problems

1. Matrix identities are commonly used in the analysis of MIMO channels. Prove the following matrix identities.

   (a) Given an $M \times N$ matrix $\mathbf{A}$, show that the matrix $\mathbf{A}\mathbf{A}^H$ is Hermitian. What does this reveal about the eigendecomposition of $\mathbf{A}\mathbf{A}^H$?

   (b) Show that $\mathbf{A}\mathbf{A}^H$ is positive semidefinite.

   (c) Show that $\mathbf{I}_M + \mathbf{A}\mathbf{A}^H$ is Hermitian positive definite.

   (d) Show that $\det[\mathbf{I}_M + \mathbf{A}\mathbf{A}^H] = \det[\mathbf{I}_N + \mathbf{A}^H\mathbf{A}]$.

2. Find the SVD of the following matrix:

$$\mathbf{H} = \begin{bmatrix} .7 & .6 & .2 & .4 \\ .1 & .5 & .9 & .2 \\ .3 & .6 & .9 & .1 \end{bmatrix}.$$

3. Find a 3×3 channel matrix $\mathbf{H}$ with two nonzero singular values.

4. Consider the 4×4 MIMO channels given below. What is the maximum multiplexing gain of each – that is, how many independent scalar data streams can be reliably supported?

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \qquad \mathbf{H}_2 = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix}.$$

5. The capacity of a static MIMO channel with only receiver CSI is given by $C = \sum_{i=1}^{R_{\mathbf{H}}} B \log_2(1 + \sigma_i^2 \rho / M_t)$. Show that, if the sum of singular values is bounded, then this expression is maximized when all $R_{\mathbf{H}}$ singular values are equal.

6. Consider a MIMO system with the following channel matrix:

$$\mathbf{H} = \begin{bmatrix} .1 & .3 & .4 \\ .3 & .2 & .2 \\ .1 & .3 & .7 \end{bmatrix} = \begin{bmatrix} -.5196 & -.0252 & -.8541 \\ -.3460 & -.9077 & .2372 \\ -.7812 & .4188 & .4629 \end{bmatrix} \begin{bmatrix} .9719 & 0 & 0 \\ 0 & .2619 & 0 \\ 0 & 0 & .0825 \end{bmatrix} \times \begin{bmatrix} -.2406 & -.8894 & .3886 \\ -.4727 & -.2423 & -.8472 \\ -.8477 & .3876 & .3621 \end{bmatrix}^H.$$

   Note that $\mathbf{H}$ is written in terms of its singular value decomposition (SVD) $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$.

   (a) Check if $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. You will see that the matrices $\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}^H$ do not have sufficiently large precision, so that $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ is only approximately equal to $\mathbf{H}$. This indicates the sensitivity of the SVD – in particular, the matrix $\mathbf{\Sigma}$ – to small errors in the estimate of the channel matrix $\mathbf{H}$.

(b) Based on the singular value decomposition $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, find an equivalent MIMO system consisting of three independent channels. Find the transmit precoding and the receiver combining matrices necessary to transform the original system into the equivalent system.

(c) Find the optimal power allocation $P_i\,(i = 1, 2, 3)$ across the three channels found in part (b), and find the corresponding total capacity of the equivalent system assuming $P/\sigma^2 = 20$ dB and a system bandwidth of $B = 100$ kHz.

(d) Compare the capacity in part (c) to that when the channel is unknown at the transmitter and so equal power is allocated to each antenna.

7. Use properties of the SVD to show that, for a MIMO channel that is known to the transmitter and receiver both, the general capacity expression

$$C = \max_{\mathbf{R}_\mathbf{x}:\mathrm{Tr}(\mathbf{R}_\mathbf{x})=\rho} B \log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R}_\mathbf{x}\mathbf{H}^H]$$

reduces to

$$C = \max_{\rho_i:\Sigma i \rho i \leq \rho} \sum_i B \log_2(1 + \sigma_i^2 \rho_i)$$

for singular values $\{\sigma_i\}$ and SNR ρ.

8. For the 4×4 MIMO channels given below, find their capacity assuming both transmitter and receiver know the channel, whose SNR $\rho = 10$ dB and bandwidth $B = 10$ MHz:

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \qquad \mathbf{H}_2 = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix}.$$

9. Assume a ZMCSCG MIMO system with channel matrix $\mathbf{H}$ corresponding to $M_t = M_r = M$ transmit and receive antennas. Using the law of large numbers, show that

$$\lim_{M\to\infty} \frac{1}{M}\mathbf{H}\mathbf{H}^H = \mathbf{I}_M.$$

Then use this to show that

$$\lim_{M\to\infty} B \log_2 \det\left[\mathbf{I}_M + \frac{\rho}{M}\mathbf{H}\mathbf{H}^H\right] = MB \log_2(1 + \rho).$$

10. Plot the ergodic capacities for a ZMCSCG MIMO channel with SNR $0 \leq \rho \leq 30$ dB and $B = 1$ MHz for the following MIMO dimensions: (a) $M_t = M_r = 1$; (b) $M_t = 2$, $M_r = 1$; (c) $M_t = M_r = 2$; (d) $M_t = 2$, $M_r = 3$; (e) $M_t = M_r = 3$. Verify that, at high SNRs, capacity grows linearly as $M = \min(M_t, M_r)$.

11. Plot the outage capacities for $B = 1$ MHz and an outage probability $P_{\mathrm{out}} = .01$ for a ZMCSCG MIMO channel with SNR $0 \leq \rho \leq 30$ dB for the following MIMO dimensions: (a) $M_t = M_r = 1$; (b) $M_t = 2$, $M_r = 1$; (c) $M_t = M_r = 2$; (d) $M_t = 2$, $M_r = 3$; (e) $M_t = M_r = 3$. Verify that, at high SNRs, capacity grows linearly as $M = \min(M_t, M_r)$.

12. Show that if the noise vector $\mathbf{n} = (n_1, \ldots, n_{M_r})$ has i.i.d. elements then, for $\|\mathbf{u}\| = 1$, the statistics of $\mathbf{u}^H\mathbf{n}$ are the same as the statistics for each of these elements.

13. Consider a MIMO system where the channel gain matrix $\mathbf{H}$ is known at the transmitter and receiver. Show that if transmit and receive antennas are used for diversity then the optimal weights at the transmitter and receiver lead to an SNR of $\gamma = \lambda_{\max}\rho$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{H}\mathbf{H}^H$. Also show that the leading eigenvectors of $\mathbf{H}^H\mathbf{H}$ and $\mathbf{H}\mathbf{H}^H$ form the optimal transmit precoding and receiver combining vectors respectively.

14. Consider a 4x4 MIMO system where the channel gain matrix $\mathbf{H}$ is known at the transmitter and receiver. Assume the channel has free space signal propagation between each transmit and receive antenna (i.e. there are no reflectors or scatterers in the environment). The antenna elements are collinear and the distance between antenna elements l at the receiver (or transmitter) is much smaller than the distance d between the antenna and the receiver. What is the approximate rank of the channel matrix $\mathbf{H}$ where singular values that are smaller than a threshold can be approximated to zero.

15. Consider a channel with channel gain matrix

$$\mathbf{H} = \begin{bmatrix} .1 & .5 & .9 \\ .3 & .2 & .6 \\ .1 & .3 & .7 \end{bmatrix}.$$

Assuming $\rho = 10$ dB, find the output SNR when beamforming is used on the channel with equal weights on each transmit antenna and optimal weighting at the receiver. Compare with the SNR under beamforming with optimal weights at both the transmitter and receiver.

16. Consider an 8×4 MIMO system, and assume a coding scheme that can achieve the rate–diversity trade-off $d(r) = (M_t - r)(M_r - r)$.

    (a) Find the maximum multiplexing rate for this channel, given a required $P_e = \rho^{-d} \leq 10^{-3}$ and assuming that $\rho = 10$ dB.

    (b) Given the r derived in part (a), what is the resulting P_e?

17. Find the capacity of a SIMO channel with channel gain vector $\mathbf{h}$ = [.1 .4 .75 .9], optimal receiver weighting, $\rho = 10$ dB, and $B = 10$ MHz.

18. Consider a 2×2 MIMO system with channel gain matrix $\mathbf{H}$ given by

$$\mathbf{H} = \begin{bmatrix} .3 & .5 \\ .7 & .2 \end{bmatrix}.$$

Assume that $\mathbf{H}$ is known at both transmitter and receiver and that there is a total transmit power of $P = 10$ mW across the two transmit antennas, AWGN with $N_0 = 10^{-9}$ W/Hz at each receive antenna, and bandwidth $B = 100$ kHz.

    (a) Find the SVD for $\mathbf{H}$.

    (b) Find the capacity of this channel.

(c) Assume that transmit precoding and receiver combining have been used to transform this channel into two parallel independent channels with a total power constraint P. Find the maximum data rate that can be transmitted over this parallel set assuming MQAM modulation on each channel, with optimal power adaptation across the channels subject to power constraint P. Assume a target BER of 10^{-3} on each channel and that the BER is bounded: $P_b \leq .2e^{-1.5\gamma/(M-1)}$; assume also that the constellation size of the MQAM is unrestricted.

(d) Suppose now that the antennas at the transmitter and receiver are all used for diversity (with optimal weighting at the transmitter and receiver) to maximize the SNR of the combiner output. Find the SNR of the combiner output as well as the BER of a BPSK modulated signal transmitted over this diversity system. Compare the data rate and BER of this BPSK signaling with diversity (assuming $B = 1/T_b$) to the rate and BER from part (c).

(e) Comment on the diversity–multiplexing trade-offs between the systems in parts (c) and (d).

19. Consider an $M \times M$ MIMO channel with ZMCSCG channel gains.

(a) Plot the ergodic capacity of this channel for $M = 1$ and $M = 4$ with $0 \leq \rho \leq 20$ dB and $B = 1$ MHz, assuming that both transmitter and receiver have CSI.

(b) Repeat part (a) assuming that only the receiver has CSI.

20. Find the outage capacity for a 4×4 MIMO channel with ZMCSCG elements at 10% outage for $\rho = 10$ dB and $B = 1$ MHz.

21. Plot the cdf of capacity for an $M \times M$ MIMO channel with $\rho = 10$ dB and $B = 1$ MHz, assuming no transmitter knowledge for $M = 4, 6, 8$. What happens as M increases? What are the implications of this behavior for a practical system design?

22. Assume a ZMCSCG MIMO system with channel matrix $\mathbf{H} \in \mathbf{R}^{M_r \times M_t}$ corresponding to M_t transmit and M_r receive antennas ($\mathrm{E}[|H_{i,j}|^2] = 1$ for all i and j). Show using the law of large numbers that

$$\lim_{M_r \to \infty} \frac{1}{M_r} \mathbf{H}^H \mathbf{H} = \mathbf{I}_{M_t}.$$

Then use this to show that

$$\lim_{M_r \to \infty} B \log_2 \det \left(\mathbf{I}_{M_t} + \frac{\rho}{M_r} \mathbf{H}^H \mathbf{H} \right) = M_t B \log_2(1 + \rho).$$

23. An 8x8 MIMO system has a channel matrix $\mathbf{H}$ with singular values $\sigma_1 = 2, \sigma_2 = 1, \sigma_3 = 0.8, \sigma_4 = 0.3, \sigma_i = 0$ where $i = 5, 6, 7, 8$. Assume the total transmit power is $\rho = 10$ dBm, the noise power at each receive antenna is 0 dBm, and the system bandwidth is B = 10 MHz. Assume in this high-tech system that both transmitter and receiver know the channel matrix.

(a) What is the rank of $\mathbf{H}$ and the SNR γ_i associated with each of its spatial dimensions assuming the full transmit power is allocated to it?

(b) Find the Shannon capacity of the system assuming optimal adaptation of power and rate across spatial dimensions.

(c) If the transmitter does not do power adaptation, it assigns equal power to all transmit antennas resulting in $\mathbf{R}_x = \frac{\rho}{8}\mathbf{I}$. Show that this implies that power is equally allocated over all 8 spatial dimensions. What is the capacity of the system in this case assuming that the rate is optimally adapted?

(d) What is the capacity when only the signal dimension with the largest SNR is used, i.e. all power is allocated to the spatial dimension with the largest singular value (beamforming).

(e) Find the maximum data rate that can be sent using adaptive modulation assuming power is optimally allocated across spatial dimensions. Compare with the maximum data rate in beamforming, where full power is allocated to the single spatial dimension associated with beamforming. In both cases assume a target BER of $P_b = 10^{-3}$ with unrestricted modulation (i.e. M can take any value). Assume a symbol time $T = 1/B$.

24. In this question we look into the diversity order of the maximum singular value $\sigma_{\max}$. Consider an $m \times n$ matrix H with i.i.d. entries distributed as Rayleigh fading variables with power p each.

(a) Use the definition of the Frobenius norm to establish an upper bound on the square of the maximum singular value of H in terms of the elements of H.

(b) Use the fact that the maximum singular value is greater than all the singular values to come up with a lower bound for $\sigma_{\max}$.

Hint: This is just a scaled version of the upper bound.

(c) What is the diversity order of the Frobenius norm? Use this result and parts (a) and (b) to write down the diversity order for the maximum singular value. For any norm $\|\cdot\|$, we define its diversity order to be

$$\lim_{p\to\infty} -\frac{\log(\text{Prob}(\|H\| < 1))}{\log(p)} \ .$$

Hint: Use results from the maximal ratio combining of Rayleigh fading random variables.

25. We are given the channel matrix $\mathbf{H}$, and a received vector $\mathbf{y}$, seen below.

$$\mathbf{H} = \begin{bmatrix} 0.3 & 0.5 & 0.5 \\ 0.7 & 0.4 & 0.4 \\ 0.4 & 0.5 & 0.6 \\ 0.2 & 0.6 & 0.7 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0.3 + 0.2i \\ 0.3 - 0.5i \\ -0.7 + 0..4i \\ -0.3 - 0.1i \end{bmatrix}$$

Assume that the transmitter used all the antennas for multiplexing gain, and transmitted 3 QPSK symbols, where the constellation points for {(00),(01),(11),(10)} correspond to {(1+i),(-1+i),(-1-i),(1- i)} respectively. Estimate the transmitted sequence assuming the following MIMO receivers:

(a) Maximum likelihood receiver

(b) Zero-forcing receiver

(c) Linear MMSE receiver (assume $N_0/E = 0.1$)

26. Consider a MIMO channel matrix $\mathbf{H}$ and a sphere decoder for decoding BPSK transmissions $\mathbf{x} \in \{-1, +1\}^n$ on the channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\nu},$$

where ν is additive white Gaussian noise. Let the decoding radius be r, i.e., the decoder solves the following problem

$$\hat{\mathbf{x}} = \text{argmin}_{\mathbf{x}\in\{-1,1\}^n : \|\mathbf{y}-\mathbf{H}\mathbf{x}\| \le r} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2.$$

(a) For what values of r will the performance of this decoder be the same as that of the maximum likelihood decoder ?

(b) For a finite value $r > 0$, we have that the sphere decoder returns $\hat{\mathbf{x}}_{\text{SD}}$. Is this the maximum likelihood estimate ?

(c) A digital communication system with one transmit and one receive antennas has maximum- likelihood (ML) receiver complexity equal to the constellation size since that is the number of possible symbols that the receiver must search over. Find the complexity of an ML MIMO receiver assuming an 8x4 MIMO system with 16 QAM modulation where the matrix is full rank and all spatial dimensions are utilized. Note that in contrast, the average complexity of sphere decoders is polynomial in the dimension when SNR is high.

27. **Sphere Decoder.** In this problem, we will take a closer look at the sphere decoder. Consider that $\mathbf{H}$ is an $n_r \times n_t$ channel gain matrix and the $n_r \geq n_t$. For this problem we will assume that $\mathbf{H}$ is full rank. The channel gain matrix can be decomposed using the QR-decomposition as

$$\mathbf{H} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix},$$

where the columns of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are all orthonormal and $\mathbf{R}$ is upper-triangular. Denote $\mathbf{y}$ as the n_r dimensional received vector and $\mathbf{x}$ as the n_t dimensional transmitted vector which consists of BPSK symbols.

(a) State the dimensions of $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{R}$, and $\mathbf{0}$.

(b) Suppose that at the receiver we are considering $\hat{\mathbf{x}}$ the transmitted signal. Show that if $\mathbf{H}\hat{\mathbf{x}}$ is within distance d of $\mathbf{y}$, then the following relation will hold

$$d^2 - \left\|\mathbf{Q}_2^H \mathbf{y}\right\|_2^2 \geq \left\|\mathbf{Q}_1^H \mathbf{y} - \mathbf{R}\hat{\mathbf{x}}\right\|_2^2.$$

(c) Denote $\tilde{\mathbf{y}} = \mathbf{Q}_1^H \mathbf{y}$. Show that the previous condition can be written as

$$d^2 - \left\|\mathbf{Q}_2^H \mathbf{y}\right\|_2 \geq \sum_{i=n_t}^{1} \left(\tilde{y}_i - \sum_{j=i}^{n_t} R_{i,j} \hat{x}_j \right)^2. \qquad (*)$$

Write the elements of the summation on the right-hand side of this expression corresponding to $i = n_t$ and $i = n_{t-1}$. In general, how many elements of $\hat{\mathbf{x}}$ are contained in the ith term of this summation?

(d) We have previously seen that, for BPSK, the ML decoder must consider all 2^{n_t} possible transmitted vectors. The relation of $(*)$ gives us an intelligent way to search this space for the most-likely transmitted vector. Consider the tree given in Figure 10.14. Each root node on the bottom of this tree corresponds to a possible $\hat{\mathbf{x}}$, where each coordinate $\hat{x}_i$ is given by the parent node contained at layer i of the tree. That is, all root nodes that are children of the node labeled '$+1$' at top layer labeled $\hat{x}_{n_t}$ have $\hat{x}_{n_t} = +1$. Describe how we can use the results of the last section to intelligently search this tree in order to find values of $\mathbf{H}\hat{\mathbf{x}}$ that are close to $\mathbf{y}$ without evaluating $\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|_2$ for all possible $\hat{\mathbf{x}}$.

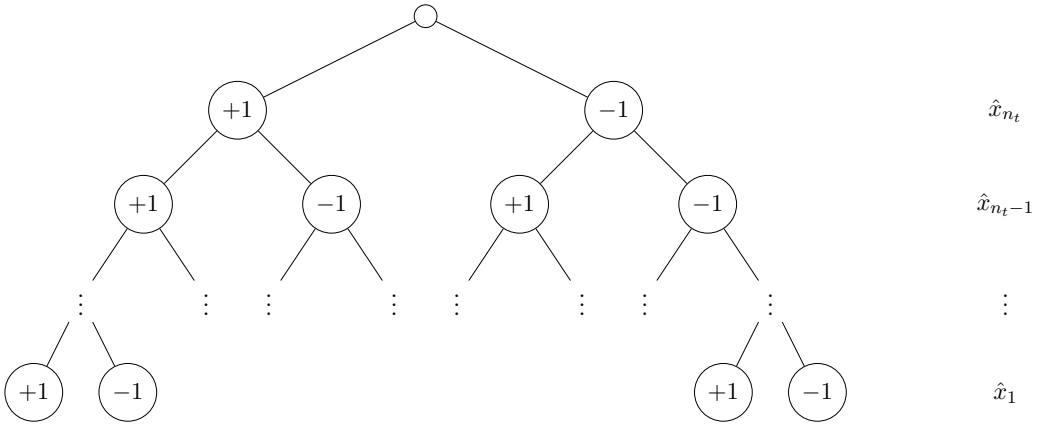


Figure 10.14: Search Tree for Problem 27

Bibliography

[1] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Sel. Areas Commun*., pp. 871–8, June 1987.

[2] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," U.S. Patent 5 345 599, 1994.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multi-element antennas," *Bell System Tech. J*., pp. 41–59, Autumn 1996.

[4] G. J. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun*., pp. 311–35, March 1998.

[5] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecommun*., pp. 585–96, November 1999.

[6] R. W. Heath and A. Lozano, *Foundations of MIMO Communications*, Cambridge University Press, 2018.

[7] J. R. Hampton, *Introduction to MIMO Communications*, Cambridge University Press, 2014.

[8] E. Biglieri, A. R. Calderbank, A. Constantinides, A.J. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications,* Cambridge University Press, 2007.

[9] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

[10] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86-105, Sept. 2007.

[11] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic and R. W. Heath, Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 122-131, December 2014.

[12] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *Bell System Tech. J*., pp. 745–78, May/June 1974.

[13] J. Salz and A. D. Wyner, "On data transmission over cross coupled multi-input, multi-output linear channels with applications to mobile radio," AT&T Bell Labs Internal Tech. Memo, 1990.

[14] B. Tsybakov, "The capacity of a memoryless Gaussian vector channel," *Prob. Inform. Trans*., 1(1), pp. 18–29, 1965.

[15] J. L. Holsinger, "Digital communication over fixed time-continuous channels with memory, with special application to telephone channels," MIT Res. Lab Elec. Tech. Rep. 430, 1964.

[16] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inform. Theory*, pp. 2636–47, October 2003.

[17] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inform. Theory*, 1(1), pp. 1–182, 2004.

[18] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186-195, February 2014.

[19] T. L. Marzetta and E. G Larsson, *Introduction to Massive MIMO*, Cambridge University Press, 2016.

[20] V. L. Girko, "A refinement of the central limit theorem for random determinants," *Theory Probab. Appl.*, 42(1), pp. 121–9, 1998.

[21] A. Grant, "Rayleigh fading multiple-antenna channels," *J. Appl. Signal Proc.*, Special Issue on Space-Time Coding (Part I), pp. 316–29, March 2002.

[22] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," *Proc. IEEE Internat. Conf. Commun.*, pp. 406–10, April 2002.

[23] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, pp. 622–40, March 1999.

[24] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," *Proc. Allerton Conf. Commun., Control, Comput.*, pp. 885–94, October 2002.

[25] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, pp. 637–50, March 2002.

[26] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inform. Theory*, pp. 2527–44, October 2003.

[27] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inform. Theory*, pp. 2545–61, October 2003.

[28] S. A. Jafar and A. J. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, pp. 1165–75, July 2004.

[29] A. Narula, M. Lopez, M. Trott, and G. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, pp. 1423–36, October 1998.

[30] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *Proc. IEEE Internat. Sympos. Inform. Theory*, pp. 357–66, June 2000.

[31] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, pp. 1543–53, September 2004.

[32] B. Hochwald and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, pp. 1893–1909, September 2004.

[33] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun*., pp. 684–701, June 2003.

[34] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, pp. 139–57, January 1999.

[35] S. A. Jafar and A. J. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun*., pp. 990–7, May 2005.

[36] L. Zheng and D. N. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multi-antenna channel," *IEEE Trans. Inform. Theory*, pp. 359–83, February 2002.

[37] R. Etkin and D. Tse, "Degrees of freedom in underspread MIMO fading channels," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 323, July 2003.

[38] A. Lapidoth and S. Moser, "On the fading number of multi-antenna systems over flat fading channels with memory and incomplete side information," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 478, July 2002.

[39] T. Koch and A. Lapidoth, "The fading number and degrees of freedom in non-coherent MIMO fading channels: A peace pipe," *Proc. IEEE Internat. Sympos. Inform. Theory*, September 2005.

[40] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications: Trends in Single- and Multi-user Systems*, Prentice-Hall, New York, 2001.

[41] A. Molisch, M. Win, and J. H. Winters, "Reduced-complexity transmit/receive-diversity systems," *IEEE Trans. Signal Proc*., pp. 2729–38, November 2003.

[42] H. Gamal, G. Caire, and M. Damon, "Lattice coding and decoding achieve the optimal diversity-multiplexing trade-off of MIMO channels," *IEEE Trans. Inform. Theory*, pp. 968–85, June 2004.

[43] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, pp. 744–65, March 1998.

[44] H. Yao and G. Wornell, "Structured space-time block codes with optimal diversity-multiplexing tradeoff and minimum delay," *Proc. IEEE Globecom Conf*., pp. 1941–5, December 2003.

[45] L. Zheng and D. N. Tse, "Diversity and multiplexing: A fundamental trade-off in multiple antenna channels," *IEEE Trans. Inform. Theory*, pp. 1073–96, May 2003.

[46] M. O. Damen, H. El Gamal, and N. C. Beaulieu, "Linear threaded algebraic space-time constellations," *IEEE Trans. Inform. Theory*, pp. 2372–88, October 2003.

[47] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, pp. 1097–1119, May 2003.

[48] R. W. Heath, Jr., and A. J. Paulraj, "Switching between multiplexing and diversity based on constellation distance," *Proc. Allerton Conf. Commun., Control, Comput*., pp. 212–21, October 2000.

[49] R. W. Heath, Jr., and D. J. Love, "Multi-mode antenna selection for spatial multiplexing with linear receivers," *IEEE Trans. Signal Proc*, pp. 3042–56, August 2005.

[50] V. Jungnickel, T. Haustein, V. Pohl, and C. Von Helmolt, "Link adaptation in a multi-antenna system," *Proc. IEEE Veh. Tech. Conf.*, pp. 862–6, April 2003.

[51] J.-C. Guey, M. P. Fitz, M. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Commun.*, pp. 527–37, April 1999.

[52] V. Tarokh, A. Naguib, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, pp. 199–207, February 1999.

[53] S. Baro, G. Bauch, and A. Hansman, "Improved codes for space-time trellis coded modulation," *IEEE Commun. Lett.*, pp. 20–2, January 2000.

[54] J. Grimm, M. Fitz, and J. Korgmeier, "Further results in space-time coding for Rayleigh fading," *Proc. Allerton Conf. Commun., Control, Comput.*, pp. 391–400, September 1998.

[55] H. Gamal and A. Hammons, "On the design of algebraic space-time codes for MIMO block-fading channels," *IEEE Trans. Inform. Theory*, pp. 151–63, January 2003.

[56] A. Naguib, N. Seshadri, and A. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Proc. Mag.*, pp. 76–92, May 2000.

[57] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, pp. 1456–67, July 1999.

[58] V. Gulati and K. R. Narayanan, "Concatenated codes for fading channels based on recursive space-time trellis codes," *IEEE Trans. Wireless Commun.*, pp. 118–28, January 2003.

[59] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Full-rate space-time codes," *IEEE J. Sel. Areas Commun.*, pp. 969–80, May 2001.

[60] K. R. Narayanan, "Turbo decoding of concatenated space-time codes," *Proc. Allerton Conf. Commun., Control, Comput.*, pp. 217–26, September 1999.

[61] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, pp. 543–64, March 2000.

[62] D. Gesbert, M. Shafi, D.-S. Shiu, P. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, pp. 281–302, April 2003.

[63] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.

[64] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel," *Proc. URSI Internat. Sympos. Signal Syst. Elec.*, pp. 295–300, October 1998.

[65] G. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Sel. Areas Commun.*, pp. 1841–52, November 1999.

[66] G. Bauch and A. Naguib, "Map equalization of space-time coded signals over frequency selective channels," *Proc. IEEE Wireless Commun. Network Conf.*, pp. 261–5, September 1999.

[67] C. Fragouli, N. Al-Dhahir, and S. Diggavi, "Pre-filtered space-time M-BCJR equalizer for frequency selective channels," *IEEE Trans. Commun*., pp. 742–53, May 2002.

[68] A. Naguib, "Equalization of transmit diversity space-time coded signals," *Proc. IEEE Globecom Conf*., pp. 1077–82, December 2000.

[69] J. Winters, "Smart antennas for wireless systems," *IEEE Pers. Commun. Mag*., pp. 23–7, February 1998.

Chapter 11

Equalization

We have seen in Chapter 6 that delay spread causes intersymbol interference (ISI), which can cause an irreducible error floor when the modulation symbol time is smaller than the channel delay spread. The delay spread for most high-speed wireless systems, whether operating indoors or outdoors, generally satisfies $\sigma_{T_m} >> T_s$. Signal processing provides a powerful mechanism to counteract ISI. In a broad sense, equalization defines any signal processing technique used at the receiver to alleviate the ISI problem caused by delay spread. Signal processing can also be used at the transmitter to make the signal less susceptible to delay spread: spread-spectrum and multicarrier modulation fall in this category of transmitter signal processing techniques. In this chapter we focus on equalization; multicarrier modulation and spread spectrum are the topics of Chapters 12 and 13, respectively. Mitigating the impact of delay spread is one of the most challenging hurdles for high-speed wireless data systems.

Equalizer design must typically balance ISI mitigation with noise enhancement, since both the signal and the noise pass through the equalizer, which can increase the noise power. Nonlinear equalizers suffer less from noise enhancement than linear equalizers but typically entail higher complexity, as discussed in more detail below. Moreover, equalizers require an estimate of the channel impulse or frequency response to mitigate the resulting ISI. Since the wireless channel varies over time, the equalizer must learn the frequency or impulse response of the channel (training) and then update its estimate of the frequency response as the channel changes (tracking). The process of equalizer training and tracking is often referred to as adaptive equalization, since the equalizer adapts to the changing channel. Equalizer training and tracking can be quite difficult if the channel is changing rapidly. In this chapter we will discuss the various issues associated with equalizer design, including balancing ISI mitigation with noise enhancement, linear and nonlinear equalizer design and properties, and the process of equalizer training and tracking.

An equalizer can be implemented at baseband, the carrier frequency, or an intermediate frequency. Most equalizers are implemented digitally after A/D conversion because such filters are small, cheap, easily tunable, and very power efficient. This chapter mainly focuses on digital equalizer implementations, although in the next section we will illustrate noise enhancement using an analog equalizer for simplicity.

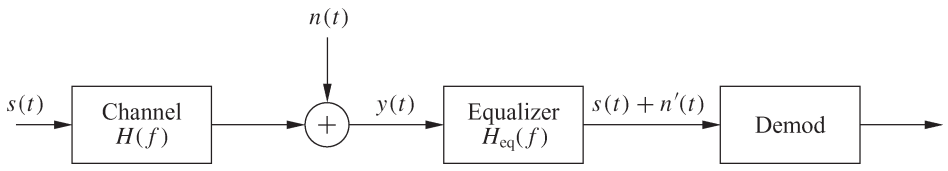


Figure 11.1: Analog equalizer illustrating noise enhancement.

11.1 ISI Mitigation and Noise Enhancement

The goal of equalization is to mitigate the effects of ISI. However, this goal must be balanced so that, in the process of removing ISI, the noise power in the received signal is not enhanced. A simple analog equalizer, shown in Figure 11.1, illustrates the pitfalls of removing ISI without considering this effect on noise. Consider a signal $s(t)$ that is passed through a channel with frequency response $H(f)$. At the receiver front end, white Gaussian noise $n(t)$ is added to the signal and so the signal input to the receiver is $Y(f) = S(f)H(f) + N(f)$, where $N(f)$ is white noise with power spectral density (PSD) $N_0/2$. If the bandwidth of $s(t)$ is B, then the noise power within the signal bandwidth of interest is N_0B. Suppose we wish to equalize the received signal so as to completely remove the ISI introduced by the channel. This is easily done by introducing an analog equalizer in the receiver that is defined by

$$H_{\text{eq}}(f) = 1/H(f). \tag{11.1}$$

The received signal $Y(f)$ after passing through this equalizer becomes

$$[S(f)H(f) + N(f)]H_{\text{eq}}(f) = S(f) + N'(f), \tag{11.2}$$

where $N'(f)$ is colored Gaussian noise with power spectral density $.5N_0/|H(f)|^2$. Thus, all ISI has been removed from the transmitted signal $S(f)$.

However, if $H(f)$ has a spectral null ($H(f_0) = 0$ for some f_0) at any frequency within the bandwidth of $s(t)$, then the power of the noise $N'(f)$ is infinite. Even without a spectral null, if some frequencies in $H(f)$ are greatly attenuated then the equalizer $H_{\text{eq}}(f) = 1/H(f)$ will greatly enhance the noise power at those frequencies. In this case, even though the ISI effects are removed, the equalized system will perform poorly because of its greatly reduced SNR. Thus, the goal of equalization is to balance mitigation of ISI with maximizing the SNR of the postequalization signal. In general, linear digital equalizers work by approximately inverting the channel frequency response and thus have the most noise enhancement. Nonlinear equalizers do not invert the channel frequency response, so they tend to suffer much less from noise enhancement. In the next section we give an overview of the different types of linear and nonlinear equalizers, their structures, and the algorithms used for updating their tap coefficients in equalizer training and tracking.

Example 11.1:
Consider a channel with impulse response $H(f) = 1/\sqrt{|f|}$ for $|f| < B$, where B is the channel bandwidth. Given noise PSD $N_0/2$, what is the noise power for channel bandwidth $B = 30$ kHz with and without an equalizer that inverts the channel?
Solution: Without equalization, the noise power is just $N_0B = 3N_0 \cdot 10^4$. With equalization, the noise PSD is $.5N_0|H_{\text{eq}}(f)|^2 = .5N_0/|H(f)|^2 = .5|f|N_0$ for $|f| < B$. So the noise power is

$$.5N_0 \int_{-B}^{B} |f| \, df = .5N_0B^2 = 4.5N_0 \cdot 10^8, \tag{11.3}$$

an increase in noise power of more than four orders of magnitude!

11.2 Equalizer Types and Structures

Equalization techniques fall into two broad categories: linear and nonlinear. The linear techniques, which perform equalization by passing the received signal through a linear filter, are generally the simplest to implement and to understand conceptually. The most common of these techniques are zero-forcing equalization (ZFE) and

minimum-mean-square-error equalization (MMSE). The ZFE can completely remove ISI at the expense of noise enhancement, whereas the MMSE can partially (rather than fully) remove ISI to reduce the noise enhancement relative to the ZFE. Linear equalization techniques, including both the ZFE and MMSE, typically suffer from more noise enhancement than nonlinear equalizers and hence nonlinear techniques are preferred in high-performance systems. Among nonlinear equalization techniques, decision-feedback equalization (DFE) is the most common because it is fairly simple to implement and usually performs well. However, on channels with low SNR, the DFE suffers from error propagation when bits are decoded in error, leading to poor performance. The optimal equalization technique for a known channel is maximum likelihood sequence estimation (MLSE), which is typically implemented using the Viterbi algorithm. The complexity of this technique grows exponentially with the length of the delay spread relative to the symbol time; this complexity precludes the use of MLSE in many high-speed data systems. However, the performance of the MLSE is often used as an upper bound on performance for other equalization techniques. Turbo equalization is an alternative nonlinear techniques to the DFE and MLSE that has shown better complexity-performance tradeoffs under certain conditions. Deep learning equalization uses training data to set the parameters of a deep neural network, which is then used to estimate the received data sequence. The neural networks in deep learning equalizers do not explicitly use channel state information or a mathematical model for the channel, in contrast to the other forms of equalization described in this chapter which either assume the channel model $H(f)$ is known or estimate it through training and tracking.

Equalizers can also be categorized as symbol-by-symbol (SBS) or sequence estimators (SEs). SBS equalizers remove ISI from each symbol and then detect each symbol individually. All linear equalizers as well as the DFE are SBS equalizers. Sequence estimators detect sequences of symbols, so the effect of ISI is part of the estimation process.

Equalizers are typically implemented digitally using a transversal or lattice structure. The transversal structure is a filter with $N-1$ delay elements and N taps featuring tunable complex weights. The lattice filter uses a more complex recursive structure [1]. In exchange for this increased complexity relative to transversal structures, lattice structures often have better numerical stability and convergence properties and greater flexibility in changing their length [2]. This chapter will focus on transversal structures; details on lattice structures and their performance relative to transversal structures can be found in [1, 2, 3, 4].

In addition to the equalizer type and structure, adaptive equalizers require algorithms for updating the filter tap coefficient. Many algorithms have been developed over the years for this purpose. These algorithms generally incorporate trade-offs between complexity, convergence rate, and numerical stability. These tradeoffs for the most common equalizer training and tracking algorithms are summarized in Table 11.1 below.

11.3 Folded Spectrum and ISI-Free Transmission

Figure 11.2 is a block diagram of an equivalent lowpass end-to-end system with a digital equalizer. The input symbol d_k is passed through a pulse-shaping filter $g(t)$ and then transmitted over the ISI channel with equivalent lowpass impulse response $c(t)$. We define the combined channel impulse response $h(t) \triangleq g(t) * c(t)$, and the equivalent lowpass transmitted signal is thus given by $d(t) * g(t) * c(t)$ for the train $d(t) = \sum_k d_k \delta(t - kT_s)$ of information symbols. The pulse shape $g(t)$ improves the spectral properties of the transmitted signal, as described in Section 5.5. This pulse shape is under the control of the system designer, whereas the channel $c(t)$ is introduced by nature and thus outside the designer's control.

At the receiver front end, equivalent lowpass white Gaussian noise $n(t)$ with PSD N_0 is added to the received signal for a resulting signal $w(t)$. This signal is passed through an analog matched filter $g_m^*(-t)$ to obtain the equivalent lowpass output $y(t)$, which is then sampled via an A/D converter. The purpose of the matched filter is to maximize the SNR of the signal before sampling and subsequent processing.[1] Recall from Section 5.1

[1] Although the matched filter could be more efficiently implemented digitally, the analog implementation before the sampler allows for

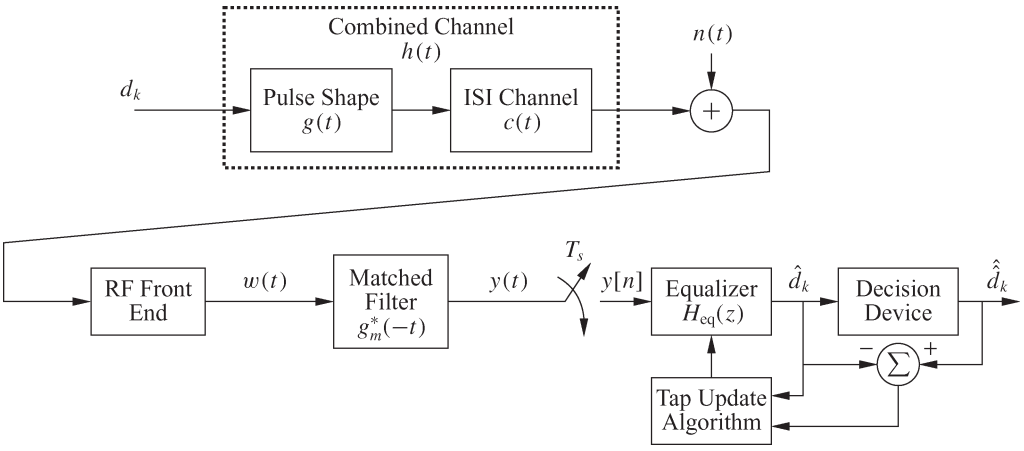


Figure 11.2: End-to-end system (equivalent lowpass representation).

that in AWGN the SNR of the received signal is maximized prior to sampling by using a filter that is matched to the pulse shape. This result indicates that, for the system shown in Figure 11.2, SNR prior to sampling is maximized by passing $w(t)$ through a filter matched to $h(t)$, so ideally we would have $g_m(t) = h(t)$. However, since the channel impulse response $c(t)$ is time varying and since analog filters are not easily tunable, it is generally not possible to have $g_m(t) = h(t)$. Thus, part of the art of equalizer design is to choose $g_m(t)$ to get good performance. Often $g_m(t)$ is matched to the pulse shape $g(t)$, which is the optimal pulse shape when $c(t) = \delta(t)$, but this design is clearly suboptimal when $c(t) \neq \delta(t)$. Not matching $g_m(t)$ to $h(t)$ can result in significant performance degradation and also makes the receiver extremely sensitive to timing error. These problems are somewhat mitigated by sampling $y(t)$ at a rate much faster than the symbol rate and then designing the equalizer for this oversampled signal. This process is called *fractionally spaced equalization* [4, Chap. 10.2.4].

The equalizer output provides an estimate of the transmitted symbol. This estimate is then passed through a decision device that rounds the equalizer output to a symbol in the alphabet of possible transmitted symbols. During training, the equalizer output is passed to the tap update algorithm to update the tap values, so that the equalizer output closely matches the known training sequence. During tracking, the round-off error associated with the symbol decision is used to adjust the equalizer coefficients.

Let $f(t)$ be defined as the composite equivalent lowpass impulse response consisting of the transmitter pulse shape, channel, and matched filter impulse responses:

$$f(t) \triangleq g(t) * c(t) * g_m^*(-t). \tag{11.4}$$

Then the matched filter output is given by

$$y(t) = d(t) * f(t) + n_g(t) = \sum d_k f(t - kT_s) + n_g(t), \tag{11.5}$$

where $n_g(t) = n(t) * g_m^*(-t)$ is the equivalent lowpass noise at the equalizer input and T_s is the symbol time. If we let $f[n] = f(nT_s)$ denote samples of $f(t)$ every T_s seconds, then sampling $y(t)$ every T_s seconds yields the discrete-time signal $y[n] = y(nT_s)$ given by

a smaller dynamic range in the sampler, which significantly reduces cost.

$$
\begin{aligned}
y[n] &= \sum_{k=-\infty}^{\infty} d_k f(nT_s - kT_s) + n_g(nT_s) \\
&= \sum_{k=-\infty}^{\infty} d_k f[n-k] + \nu[n] \\
&= d_n f[0] + \sum_{k \neq n} d_k f[n-k] + \nu[n], \qquad (11.6)
\end{aligned}
$$

where the first term in (11.6) is the desired data bit, the second term is the ISI, and the third term is the sampled noise. By (11.6) there is no ISI if $f[n-k] = 0$ for $k \neq n$ – that is, $f[k] = \delta[k]f[0]$. In this case (11.6) reduces to $y[n] = d_n f[0] + \nu[n]$.

We now show that the condition for ISI-free transmission, $f[k] = \delta[k]f[0]$, is satisfied if and only if

$$
F_\Sigma(f) \triangleq \frac{1}{T_s} \sum_{n=-\infty}^{\infty} F\left(f + \frac{n}{T_s}\right) = f[0]. \qquad (11.7)
$$

The function $F_\Sigma(f)$, which is periodic with period $1/T_s$, is often called the *folded spectrum*. If $F_\Sigma(f) = f[0]$ we say that the folded spectrum is *flat*.

To show the equivalence between ISI-free transmission and a flat folded spectrum, begin by observing that

$$
\begin{aligned}
f[k] = f(kT_s) &= \int_{-\infty}^{\infty} F(f) e^{j2\pi f k T_s} df \\
&= \sum_{n=-\infty}^{\infty} \int_{.5(2n-1)/T_s}^{.5(2n+1)/T_s} F(f) e^{j2\pi f k T_s} df \\
&= \sum_{n=-\infty}^{\infty} \int_{-.5/T_s}^{.5/T_s} F\left(f' + \frac{n}{T_s}\right) e^{j2\pi(f' + n/T_s)kT_s} df' \\
&= \int_{-.5/T_S}^{.5/T_S} e^{j2\pi f k T_s} \left[\sum_{n=-\infty}^{\infty} F\left(f + \frac{n}{T_s}\right)\right] df. \qquad (11.8)
\end{aligned}
$$

Equation (11.8) implies that the Fourier series representation of $F_\Sigma(f)$ is

$$
F_\Sigma(f) = \frac{1}{T_s} \sum_k f[k] e^{-j2\pi f k T_s}.
$$

We first demonstrate that a flat folded spectrum implies that $f[k] = \delta[k]f[0]$. Suppose (11.7) is true. Then substituting (11.7) into (11.8) yields

$$
f[k] = T_s \int_{-.5/T_S}^{.5/T_S} e^{j2\pi f k T_s} f[0] df = \frac{\sin \pi k}{\pi k} f[0] = \delta[k] f[0], \qquad (11.9)
$$

which is the desired result. We now show that $f[k] = \delta[k]f[0]$ implies a flat folded spectrum. By (11.8) and the definition of $F_\Sigma(f)$,

$$
f[k] = T_s \int_{-.5/T_S}^{.5/T_S} F_\Sigma(f) e^{j2\pi f k T_s} df. \qquad (11.10)
$$

Hence $f[k]$ is the inverse Fourier transform of $F_\Sigma(f)$.[2] Therefore, if $f[k] = \delta[k]f[0]$ then $F_\Sigma(f) = f[0]$.

Example 11.2:

Consider a channel with combined equivalent lowpass impulse response $f(t) = \text{sinc}(t/T_s)$. Find the folded spectrum and determine if this channel exhibits ISI.

Solution: The Fourier transform of $f(t)$ is

$$F(f) = T_s \text{ rect}(fT_s) = \begin{cases} T_s & |f| < .5/T_s, \\ .5T_s & |f| = .5/T_s, \\ 0 & |f| > .5/T_s. \end{cases} \tag{11.11}$$

Therefore,

$$F_\Sigma(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} F\left(f + \frac{n}{T_s}\right) = 1, \tag{11.12}$$

so the folded spectrum is flat and there is no ISI. We can also see this from the fact that

$$f(nT_s) = \text{sinc}\left(\frac{nT_s}{T_s}\right) = \text{sinc}(n) = \begin{cases} 1 & n = 0, \\ 0 & n \neq 0. \end{cases} \tag{11.13}$$

Thus, $f[k] = \delta[k]$, our equivalent condition for a flat folded spectrum and zero ISI.

Channel equalization generally takes place in the time domain, with the digital filter $H_{\text{eq}}(z)$ in Figure 11.2 operating on the sampled output signal $y[n]$. For channels with large delay spread, this can entail significant complexity and delay since the size of the equalizer is on the order of T_s/T_m. For such channels, performing equalization in the frequency domain can reduce complexity and delay over time domain processing [5]. This technique, called *frequency-domain equalization*, converts a block of N time samples $y[n], n = 1, \ldots, N$ via an N-point FFT to the frequency domain and performs equalization in that domain, then converts the signal back to the time domain via an N-point IFFT. When N is selected such that $NT_s >> T_m$, each component in the FFT output vector $Y[n], n = 1, \ldots, N$ experiences flat-fading and hence can be equalized via scalar multiplication. This phenomenon leads to the use of frequency-domain equalization for multicarrier modulation and its OFDM implementation, as will be discussed in Chapter 12.4.3.

11.4 Linear Equalizers

If $F_\Sigma(f)$ is not flat, we can use the equalizer $H_{\text{eq}}(z)$ in Figure 11.2 to reduce ISI. In this section we assume a linear equalizer implemented via a $2L + 1 = N$-tap transversal filter:

$$H_{\text{eq}}(z) = \sum_{i=-L}^{L} w_i z^{-i}. \tag{11.14}$$

The length of the equalizer N is typically dictated by implementation considerations, since a large N entails more complexity and delay. Causal linear equalizers have $w_i = 0$ when $i < 0$. For a given equalizer size N, the equalizer design must specify (i) the tap weights $\{w_i\}_{i=-L}^{L}$ for a given channel frequency response and (ii) the algorithm for updating these tap weights as the channel varies. Recall that our performance metric in wireless systems is

[2] This also follows from the Fourier series representation of $F_\Sigma(f)$.

probability of error (or outage probability), so for a given channel the optimal choice of equalizer coefficients would be the coefficients that minimize probability of error. Unfortunately, it is extremely difficult to optimize the $\{w_i\}$ with respect to this criterion. Since we cannot directly optimize for our desired performance metric, we must instead use an indirect optimization that balances ISI mitigation with the prevention of noise enhancement, as discussed with regard to the preceding simple analog example. We now describe two linear equalizers: the zero-forcing (ZF) equalizer and the minimum mean-square error (MMSE) equalizer. The former equalizer cancels all ISI but can lead to considerable noise enhancement. The latter technique minimizes the expected mean-squared error between the transmitted symbol and the symbol detected at the equalizer output, thereby providing a better balance between ISI mitigation and noise enhancement. Because of this more favorable balance, MMSE equalizers tend to have better BER performance than equalizers using the ZF algorithm.

11.4.1 Zero-Forcing (ZF) Equalizers

From (11.6), the samples $\{y_n\}$ input to the equalizer can be represented based on the discretized combined equivalent lowpass impulse response $f(t) = h(t) * g^*(-t)$ as

$$Y(z) = D(z)F(z) + N_g(z), \tag{11.15}$$

where $N_g(z)$ is the z-transform of the noise samples at the output of the matched filter $G_m^*(1/z^*)$ and

$$F(z) = H(z)G_m^*\left(\frac{1}{z^*}\right) = \sum_n f(nT_s)z^{-n}. \tag{11.16}$$

The zero-forcing equalizer removes all ISI introduced in the composite response $f(t)$. From (11.15) we see that the equalizer to accomplish this is given by

$$H_{\text{ZF}}(z) = \frac{1}{F(z)}. \tag{11.17}$$

This is the discrete-time equivalent lowpass equalizer of the analog equalizer (11.1) described before, and it suffers from the same noise enhancement properties. Specifically, the power spectrum $N(z)$ of the noise samples at the equalizer output is given by

$$\begin{aligned} N(z) &= N_0|G_m^*(1/z^*)|^2|H_{\text{ZF}}(z)|^2 = \frac{N_0|G_m^*(1/z^*)|^2}{|F(z)|^2} \\ &= \frac{N_0|G_m^*(1/z^*)|^2}{|H(z)|^2|G_m^*(1/z^*)|^2} = \frac{N_0}{|H(z)|^2}. \end{aligned} \tag{11.18}$$

We see from (11.18) that if the channel $H(z)$ is sharply attenuated at any frequency within the signal bandwidth of interest – as is common on frequency-selective fading channels – the noise power will be significantly increased. This motivates an equalizer design that better optimizes between ISI mitigation and noise enhancement. One such equalizer is the MMSE equalizer, described in the next section.

The ZF equalizer defined by $H_{\text{ZF}}(z) = 1/F(z)$ may not be implementable as a finite-impulse response (FIR) filter. Specifically, it may not be possible to find a finite set of coefficients $w_{-L}, \ldots, w_L$ such that

$$w_{-L}z^L + \cdots + w_L z^{-L} = \frac{1}{F(z)}. \tag{11.19}$$

In this case we find the set of coefficients $\{w_i\}$ that best approximates the zero-forcing equalizer. Note that this is not straightforward because the approximation must be valid for all values of z. There are many ways we

can make this approximation. One technique is to represent $H_{\rm ZF}(z)$ as an infinite–impulse response (IIR) filter, $1/F(z) = \sum_{i=-\infty}^{\infty} c_i z^{-i}$, and then set $w_i = c_i$. It can be shown that this minimizes

$$\left| \frac{1}{F(z)} - (w_{-L} z^L + \cdots + w_L z^{-L}) \right|^2 \tag{11.20}$$

at $z = e^{j\omega}$. Alternatively, the tap weights can be set to minimize the peak distortion (worstcase ISI). Finding the tap weights to minimize peak distortion is a convex optimization problem that can be solved by standard techniques – for example, the method of steepest descent [4].

Example 11.3:

Consider a channel with impulse response

$$h(t) = \begin{cases} e^{-t/\tau} & t \geq 0, \\ 0 & \text{else.} \end{cases} \tag{11.21}$$

Find a two-tap ZF equalizer for this channel.
Solution: We have

$$h[n] = 1 + e^{-T_s/\tau}\delta[n-1] + e^{-2T_s/\tau}\delta[n-2] + \cdots . \tag{11.22}$$

Thus,

$$\begin{aligned} H(z) &= 1 + e^{-T_s/\tau} z^{-1} + e^{-2T_s/\tau} z^{-2} + e^{-3T_s/\tau} z^{-3} + \cdots \\ &= \sum_{n=0}^{\infty} (e^{-T_s/\tau} z^{-1})^n = \frac{z}{z - e^{-T_s/\tau}}. \end{aligned} \tag{11.23}$$

So $H_{\rm eq}(z) = 1/H(z) = 1 - e^{-T_s/\tau} z^{-1}$. The two-tap ZF equalizer therefore has tap weight coefficients $w_0 = 1$ and $w_1 = e^{-T_s/\tau}$.

If the channel is known at the transmitter, then the transmitter can *pre-equalize* the transmitted signal by passing it through a zero-forcing filter that effectively inverts the channel frequency response. Since the channel inversion occurs in the transmitter rather than the receiver, there is no noise enhancement. It is difficult to pre-equalize in a time-varying channel because the transmitter must have an accurate estimate of the channel, but this approach is practical to implement in relatively static wireline channels. One problem is that the channel inversion often increases the dynamic range of the transmitted signal, which can result in distortion or inefficiency from the amplifier. This problem has been addressed through a precoding technique called *Tomlinson–Harashima precoding* [19, 20].

11.4.2 Minimum Mean-Square Error (MMSE) Equalizers

In MMSE equalization, the goal of the equalizer design is to minimize the average mean square error (MSE) between the transmitted symbol d_k and its estimate $\hat{d}_k$ at the output of the equalizer. In other words, the $\{w_i\}$ are chosen to minimize $\mathbf{E}[d_k - \hat{d}_k]^2$. Since the MMSE equalizer is linear, its output $\hat{d}_k$ is a linear combination of the input samples $y[k]$:

$$\hat{d}_k = \sum_{i=-L}^{L} w_i y[k-i]. \tag{11.24}$$

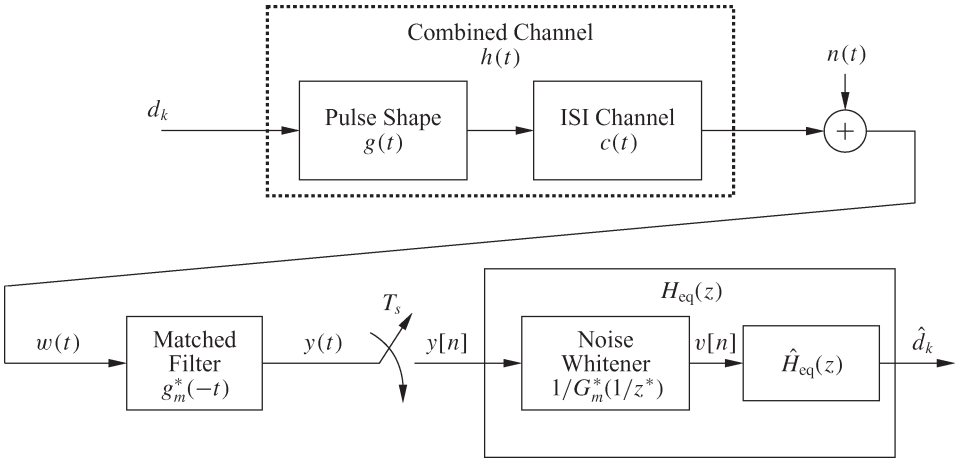


Figure 11.3: MMSE equalizer with noise-whitening filter.

As such, finding the optimal filter coefficients $\{w_i\}$ becomes a standard problem in linear estimation. In fact, if the noise input to the equalizer is white then we have a standard Weiner filtering problem. However, because of the matched filter $g_m^*(-t)$ at the receiver front end, the noise input to the equalizer is not white; rather, it is colored with power spectrum $N_0|G_m^*(1/z^*)|^2$. Therefore, in order to apply known techniques for optimal linear estimation, we expand the filter $H_{\text{eq}}(z)$ into two components – a noise-whitening component $1/G_m^*(1/z^*)$ and an ISI-removal component $\hat{H}_{\text{eq}}(z)-$ as shown in Figure 11.3.

The purpose of the noise-whitening filter (as the name indicates) is to whiten the noise so that the noise component output from this filter has a constant power spectrum. Since the noise input to this filter has power spectrum $N_0|G_m^*(1/z^*)|^2$, the appropriate noise-whitening filter is $1/G_m^*(1/z^*)$. The noise power spectrum at the output of the noise-whitening filter is then $N_0|G_m^*(1/z^*)|^2/|G_m^*(1/z^*)|^2 = N_0$. Note that the filter $1/G_m^*(1/z^*)$ is not the only filter that will whiten the noise, and another noise-whitening filter with more desirable properties (like stability) may be chosen. It might seem odd at first to introduce the matched filter $g_m^*(-t)$ at the receiver front end only to cancel its effect in the equalizer. Recall, however, that the matched filter is meant to maximize the SNR prior to sampling. By removing the effect of this matched filter via noise whitening after sampling, we merely simplify the design of $\hat{H}_{\text{eq}}(z)$ to minimize MSE. In fact, if the noise-whitening filter does not yield optimal performance then its effect would be cancelled by the $\hat{H}_{\text{eq}}(z)$ filter design, as we shall see in the case of IIR MMSE equalizers.

We assume that the filter $\hat{H}_{\text{eq}}(z)$, with input v_n, is a linear filter with $N = 2L + 1$ taps:

$$\hat{H}_{\text{eq}}(z) = \sum_{i=-L}^{L} w_i z^{-i}. \tag{11.25}$$

Our goal is to design the filter coefficients $\{w_i\}$ so as to minimize $\mathbf{E}[d_k - \hat{d}_k]^2$. This is the same goal as for the total filter $H_{\text{eq}}(z)$ – we just added the noise-whitening filter to make solving for these coefficients simpler. Define $\mathbf{v}^T = (v[k+L], v[k+L-1], \ldots, v[k-L]) = (v_{k+L}, v_{k+L-1}, \ldots, v_{k-L})$ as the row vector of inputs to the filter $\hat{H}_{\text{eq}}(z)$ used to obtain the filter output $\hat{d}_k$ and define $\mathbf{w}^T = (w_{-L} \ldots, w_L)$ as the row vector of filter coefficients. Then

$$\hat{d}_k = \mathbf{w}^T\mathbf{v} = \mathbf{v}^T\mathbf{w}. \tag{11.26}$$

Thus, we want to minimize the mean-square error

$$J = \mathbf{E}[d_k - \hat{d}_k]^2 = \mathbf{E}[\mathbf{w}^T\mathbf{v}\mathbf{v}^H\mathbf{w}^* - 2\mathrm{Re}\{\mathbf{v}^H\mathbf{w}^*\mathbf{d}_k\} + |\mathbf{d}_k|^2]. \tag{11.27}$$

Define $\mathbf{M_v} = \mathbf{E}[\mathbf{v}\mathbf{v}^H]$ and $\mathbf{v_d}^T = \mathbf{E}[\mathbf{v}^H d_k]$. The matrix $\mathbf{M}_\mathrm{v}$ is an $N \times N$ Hermitian matrix, and $\mathbf{v_d}$ is a length-N column vector. Assume $\mathbf{E}|d_k|^2 = 1$. Then the MSE J is

$$J = \mathbf{w}^T\mathbf{M_v}\mathbf{w}^* - 2\,\mathrm{Re}\{\mathbf{v_d}^T\mathbf{w}^*\} + 1. \tag{11.28}$$

Let $\mathbf{w} = \mathbf{x} + i\mathbf{y}$, $\mathbf{M_v} = \mathbf{A} + i\mathbf{B}$, and $\mathbf{v_d} = \mathbf{a} + i\mathbf{b}$, where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{a}$, $\mathbf{b}$ are real vectors, and $\mathbf{A}$, $\mathbf{B}$ are real matrices. Recall that $\mathbf{M_v}^H = \mathbf{M_v}$, we have $\mathbf{A}^T = \mathbf{A}$, and $\mathbf{B}^T = -\mathbf{B}$. We obtain the optimal tap vector $\mathbf{w}$ by setting the gradients $\nabla_\mathbf{x} J = \nabla_\mathbf{y} J = 0$ and solving for $\mathbf{x}, \mathbf{y}$. By (11.28), we have

$$J = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{A}\mathbf{y} + \mathbf{x}^T\mathbf{B}\mathbf{y} - \mathbf{y}^T\mathbf{B}\mathbf{x} - 2\mathbf{a}^T\mathbf{x} - 2\mathbf{b}^T\mathbf{y} + 1 + i(\mathbf{x}^T\mathbf{B}\mathbf{x} + \mathbf{y}^T\mathbf{B}\mathbf{y} + \mathbf{y}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{A}\mathbf{y}). \tag{11.29}$$

Since $\mathbf{B}^T = -\mathbf{B}$, we have $\mathbf{x}^T\mathbf{B}\mathbf{x} = \mathbf{y}^T\mathbf{B}\mathbf{y} = 0$, and $\mathbf{x}^T\mathbf{B}\mathbf{y} = -\mathbf{y}^T\mathbf{B}\mathbf{x}$. With $\mathbf{A}^T = \mathbf{A}$, we have $\mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{y}$. Hence,

$$J = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{A}\mathbf{y} + \mathbf{x}^T\mathbf{B}\mathbf{y} - \mathbf{y}^T\mathbf{B}\mathbf{x} - 2\mathbf{a}^T\mathbf{x} - 2\mathbf{b}^T\mathbf{y} + 1, \tag{11.30}$$

whose gradients are given by

$$\nabla_\mathbf{x} J = 2\mathbf{x}^T\mathbf{A} - 2\mathbf{y}^T\mathbf{B} - 2\mathbf{a}^T, \tag{11.31}$$

and

$$\nabla_\mathbf{y} J = 2\mathbf{y}^T\mathbf{A} + 2\mathbf{x}^T\mathbf{B} - 2\mathbf{b}^T. \tag{11.32}$$

Setting $\nabla_\mathbf{x} J = \nabla_\mathbf{y} J = 0$ leads to

$$0 = \nabla_\mathbf{x} J + i\nabla_\mathbf{y} J = 2\mathbf{w}^T\mathbf{M_v} - 2\mathbf{v_d}^T, \tag{11.33}$$

which leads to $\mathbf{w}^T\mathbf{M_v} = \mathbf{v_d}^T$ or, equivalently, that the optimal tap weights are given by

$$\mathbf{w}_\mathrm{opt} = (\mathbf{M_v}^T)^{-1}\mathbf{v_d}. \tag{11.34}$$

Note that solving for $\mathbf{w}_\mathrm{opt}$ requires a matrix inversion with respect to the filter inputs. Thus, the complexity of this computation is quite high, typically on the order of N^2 to N^3 operations. Substituting in these optimal tap weights, we obtain the minimum mean-square error as

$$J_\mathrm{min} = 1 - \mathbf{v_d}^T\mathbf{M_v}^{-1}\mathbf{v_d}^*. \tag{11.35}$$

For an equalizer of infinite length, $\mathbf{v}^T = (v_{n+\infty},\, \ldots,\, v_n,\, \ldots,\, v_{n-\infty})$ and $\mathbf{w}^T = (w_{-\infty}, \ldots,\, w_0, \ldots, w_\infty)$. Then $\mathbf{w}^T\mathbf{M_v} = \mathbf{v_d}^T$ can be written as

$$\sum_{i=-\infty}^{\infty} w_i(f[j-i] + N_0\delta[j-i]) = g_m^*[-j], \quad -\infty \le j \le \infty \tag{11.36}$$

(see [6, Chap. 7.4]). Taking z-transforms and noting that $\hat{H}_\mathrm{eq}(z)$ is the z-transform of the filter coefficients $\mathbf{w}$ yields

$$\hat{H}_\mathrm{eq}(z)(F(z) + N_0) = G_m^*(1/z^*). \tag{11.37}$$

Solving for $\hat{H}_\mathrm{eq}(z)$, we obtain

$$\hat{H}_{\rm eq}(z) = \frac{G_m^*(1/z^*)}{F(z) + N_0}. \tag{11.38}$$

Since the MMSE equalizer consists of the noise-whitening filter $1/G_m^*(1/z^*)$ plus the ISI removal component $\hat{H}_{\rm eq}(z)$, it follows that the full MMSE equalizer (when it is not restricted to be finite length) becomes

$$H_{\rm eq}(z) = \frac{\hat{H}_{\rm eq}(z)}{G_m^*(1/z^*)} = \frac{1}{F(z) + N_0}. \tag{11.39}$$

There are three interesting things to note about this result. First of all, the ideal infinite length MMSE equalizer cancels out the noise-whitening filter. Second, this infinite-length equalizer is identical to the ZF filter except for the noise term N_0, so in the absence of noise the two equalizers are equivalent. Finally, this ideal equalizer design clearly shows a balance between inverting the channel and noise enhancement: if $F(z)$ is highly attenuated at some frequency, then the noise term N_0 in the denominator prevents the noise from being significantly enhanced by the equalizer. Yet at frequencies where the noise power spectral density N_0 is small compared to the composite channel $F(z)$, the equalizer effectively inverts $F(z)$.

For the equalizer (11.39) it can be shown [6, Chap. 7.4.1] that the minimum MSE (11.35) can be expressed in terms of the folded spectrum $F_\Sigma(f)$ as

$$J_{\rm min} = T_s \int_{-.5/T_s}^{.5/T_s} \frac{N_0}{F_\Sigma(f) + N_0} df. \tag{11.40}$$

This expression for MMSE has several interesting properties. First it can be shown that, as expected, $0 \le J_{\rm min} = \mathbf{E}[d_k - \hat{d}_k]^2 \le 1$. In addition, $J_{\rm min} = 0$ in the absence of noise ($N_0 = 0$) as long as $F_\Sigma(f) \neq 0$ within the signal bandwidth of interest. Also, as expected, $J_{\rm min} = 1$ if $N_0 = \infty$.

Example 11.4:
Find $J_{\rm min}$ when the folded spectrum $F_\Sigma(f)$ is flat, $F_\Sigma(f) = f[0]$, in the asymptotic limit of high and low SNR.

Solution: If $F_\Sigma = f[0] \triangleq f_0$ then

$$J_{\rm min} = T_s \int_{-.5/T_s}^{.5/T_s} \frac{N_0}{f_0 + N_0} df = \frac{N_0}{f_0 + N_0}. \tag{11.41}$$

For high SNR, $f_0 \gg N_0$ and so $J_{\rm min} \approx N_0/f_0 = N_0/E_s$, where E_s/N_0 is the SNR per symbol. For low SNR, $N_0 \gg f_0$ and so $J_{\rm min} = N_0/(N_0 + f_0) \approx N_0/N_0 = 1$.

11.5 Nonlinear Equalizers

11.5.1 Maximum Likelihood Sequence Estimation

Maximum-likelihood sequence estimation avoids the problem of noise enhancement because it doesn't use an equalizing filter: instead it estimates the sequence of transmitted symbols. The structure of the MLSE is the same as in Figure 11.2, except that the equalizer $H_{\rm eq}(z)$ and decision device are replaced by the MLSE algorithm. Given the combined pulse-shaping filter and channel response $h(t)$, the MLSE algorithm chooses the input sequence $\{d_k\}$ that maximizes the likelihood of the received signal $w(t)$. We now investigate this algorithm in more detail.

Using a Gram–Schmidt orthonormalization procedure, we can express $w(t)$ on a time interval $[0, LT_s]$ as

$$w(t) = \sum_{n=1}^{N} w_n \phi_n(t), \tag{11.42}$$

where $\{\phi_n(t)\}$ form a complete set of orthonormal basis functions. The number N of functions in this set is a function of the channel memory, since $w(t)$ on $[0, LT_s]$ depends on $d_0, \ldots, d_L$. With this expansion we have

$$w_n = \sum_{k=-\infty}^{\infty} d_k h_{nk} + \nu_n = \sum_{k=0}^{L} d_k h_{nk} + \nu_n, \tag{11.43}$$

where

$$h_{nk} = \int_0^{LT_s} h(t - kT_s)\phi_n^*(t)\, dt \tag{11.44}$$

and

$$\nu_n = \int_0^{LT_s} n(t)\phi_n^*(t)\, dt. \tag{11.45}$$

The ν_n are complex Gaussian random variables with mean zero and covariance .5 $\mathbf{E}[\nu_n^*\nu_m] = N_0\delta[n-m]$. Thus, $\mathbf{w}^N = (w_1, \ldots, w_N)$ has a multivariate Gaussian distribution:

$$p\left(\mathbf{w}^N | d^L, h(t)\right) = \prod_{n=1}^{N} \left[\frac{1}{\pi N_0} \exp\left[-\frac{1}{N_0} \left| w_n - \sum_{k=0}^{L} d_k h_{nk} \right|^2 \right] \right]. \tag{11.46}$$

Given a received signal $w(t)$ or, equivalently, $\mathbf{w}^N$, the MLSE decodes this as the symbol sequence d^L that maximizes the likelihood function $p(\mathbf{w}^N | d^L, h(t))$ (or the log of this function). That is, the MLSE outputs the sequence

$$\begin{aligned}
\hat{d}^L &= \arg\max[\log p(\mathbf{w}^N | d^L, h(t))] \\
&= \arg\max \left[-\sum_{n=1}^{N} \left| w_n - \sum_k d_k h_{nk} \right|^2 \right] \\
&= \arg\max \Bigg[-\sum_{n=1}^{N} |w_n|^2 + \sum_{n=1}^{N} \left(w_n^* \sum_k d_k h_{nk} + w_n \sum_k d_k^* h_{nk}^* \right) \\
&\qquad - \sum_{n=1}^{N} \left(\sum_k d_k h_{nk} \right) \left(\sum_m d_m^* h_{nm}^* \right) \Bigg] \\
&= \arg\max \left[2\ \mathrm{Re} \left\{ \sum_k d_k^* \sum_{n=1}^{N} w_n h_{nk}^* \right\} - \sum_k \sum_m d_k d_m^* \sum_{n=1}^{N} h_{nk} h_{nm}^* \right].
\end{aligned} \tag{11.47}$$

Note that

$$\sum_{n=1}^{N} w_n h_{nk}^* = \int_{-\infty}^{\infty} w(\tau) h^*(\tau - kT_s)\, d\tau = y[k] \tag{11.48}$$

and

$$\sum_{n=1}^{N} h_{nk}h_{nm}^{*} = \int_{-\infty}^{\infty} h(\tau - kT_s)h^{*}(\tau - mT_s)\, d\tau = f[k-m]. \tag{11.49}$$

Let $u(t) = h(t) * h^{*}(-t)$. Combining (11.47), (11.48), and (11.49), we have that

$$\hat{d}^{L} = \arg\max \left[2 \text{ Re} \left\{ \sum_{k} d_k^{*} y[k] \right\} - \sum_{k} \sum_{m} d_k d_m^{*} u[k-m] \right]. \tag{11.50}$$

We see from this equation that the MLSE output depends only on the sampler output $\{y[k]\}$ and the channel parameters $u[n-k] = u(nT_s - kT_s)$. Because the derivation of the MLSE is based on the channel output $w(t)$ only (prior to matched filtering), our derivation implies that the receiver matched filter in Figure 11.2 with $g_m(t) = h(t)$ is optimal for MLSE detection (typically the matched filter is optimal for detecting signals in AWGN, but this derivation shows that it is also optimal for detecting signals in the presence of ISI if MLSE is used).

The most common method to implement a MLSE is the Viterbi algorithm, described in Chapter 8.4.3 as a reduced-complexity optimal decoder for convolution codes. The same properties by which the Viterbi algorithm reduces the complexity of optimal convolutional decoding are used to reduce the complexity of the MLSE: in this setting the algorithm discards some possible transmitted symbol sequences - those that cannot be the ML sequence - before all sequences have been evaluated [4, 6, 7, 8]. While this reduces the complexity of MLSE relative to an exhaustive search, the complexity still grows exponentially with the channel delay spread. Techniques to avoid this exponential complexity in the Viterbi MLSE include bounding the number of its surviving sequences as well as reducing the number of symbols spanned by the ISI through preprocessing or decision-feedback, as described in [4, Chap. 10.4]. These reduced-complexity Viterbi equalizers have better performance-complexity trade-offs than symbol-by-symbol equalization techniques, and they achieve performance close to that of the optimal MLSE with significantly less complexity. However, the complexity is still prohibitive for some systems where complexity is a significant constraint. For such systems nonlinear equalization using decision feedback equalization, described in the next section, is typically utilized as providing better performance than linear techniques without requiring significantly higher complexity.

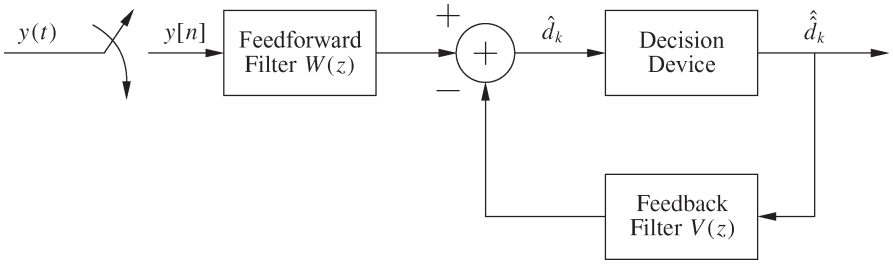


Figure 11.4: Decision-feedback equalizer structure.

11.5.2 Decision-Feedback Equalization

The DFE consists of a feedforward filter $W(z)$ with the received sequence as input (similar to the linear equalizer) followed by a feedback filter $V(z)$ with the previously detected sequence as input. The DFE structure is shown in Figure 11.4. In effect, the DFE determines the ISI contribution from the detected symbols $\left\{\hat{\hat{d}}_k\right\}$ by passing them through a feedback filter that approximates the composite channel $F(z)$ convolved with the feedforward

filter $W(z)$. The resulting ISI is then subtracted from the incoming symbols. Since the feedback filter $V(z)$ in Figure 11.4 sits in a feedback loop, it must be strictly causal or else the system is unstable. The feedback filter of the DFE does not suffer from noise enhancement because it estimates the channel frequency response rather than its inverse. For channels with deep spectral nulls, DFEs generally perform much better than linear equalizers.

Assuming that $W(z)$ has $N_1 + 1$ taps and $V(z)$ has N_2 taps, we can write the DFE output as

$$\hat{d}_k = \sum_{i=-N_1}^{0} w_i y[k-i] - \sum_{i=1}^{N_2} v_i \hat{\hat{d}}_{k-i}. \tag{11.51}$$

The typical criteria for selecting the coefficients for $W(z)$ and $V(z)$ are either zero-forcing (remove all ISI) or MMSE (minimize expected MSE between the DFE output and the original symbol). We will assume that the matched filter in the receiver is perfectly matched to the channel: $g_m(t) = h(t)$. With this assumption, in order to obtain a causal filter for $V(z)$, we require a spectral factorization of $H(z)G_m^*(1/z^*) = H(z)H^*(1/z^*)$ [4, Chap. 10.1.2]. Specifically, $H(z)H^*(1/z^*)$ is factorized by finding a causal filter $B(z)$ such that $H(z)H^*(1/z^*) = B(z)B^*(1/z^*) = \lambda^2 B_1(z)B_1^*(1/z^*)$, where $B_1(z)$ is normalized to have a leading coefficient of 1. With this factorization, when both $W(z)$ and $V(z)$ have infinite duration, it was shown by Price [9] that the optimal feedforward and feedback filters for a zero-forcing DFE are (respectively) $W(z) = 1/(\lambda^2 B_1^*(1/z^*))$ and $V(z) = 1 - B_1(z)$.

For the MMSE criterion, we wish to minimize $\mathbf{E}[d_k - \hat{d}_k]^2$. Let $b_n = b[n]$ denote the inverse z-transform of $B(z)$ given the spectral factorization $H(z)H^*(1/z^*) = B(z)B^*(1/z^*)$ just described. Then the MMSE minimization yields that the coefficients of the feedforward filter must satisfy the following set of linear equations:

$$\sum_{i=-N_1}^{0} q_{li}\hat{w}_i = b_{-l}^* \tag{11.52}$$

for $q_{li} = \sum_{j=-l}^{0} b_j^* b_{j+l-i} + N_0\delta[l-i]$ with $i = -N_1, \ldots, 0$ [4, Chap. 10.3.1]. The coefficients of the feedback filter are then determined from the feedforward coefficients by

$$v_k = -\sum_{i=-N_1}^{0} \hat{w}_i b_{k-i}. \tag{11.53}$$

These coefficients completely eliminate ISI when there are no decision errors – that is, when $\hat{\hat{d}}_k = d_k$. It was shown by Salz [10] that the resulting minimum MSE is

$$J_{\min} = \exp\left[T_s \int_{-.5/T_s}^{.5/T_s} \ln\left[\frac{N_0}{F_\Sigma(f) + N_0}\right] df\right]. \tag{11.54}$$

In general, the MMSE associated with a DFE is much lower than that of a linear equalizer, assuming the impact of feedback errors is ignored.

Decision-feedback equalizers exhibit feedback errors if $\hat{\hat{d}}_k \neq d_k$, since the ISI subtracted from the feedback path is not the true ISI corresponding to d_k. Such errors therefore propagate to later bit decisions. Moreover, this error propagation cannot be improved through channel coding, since the feedback path operates on coded channel symbols before decoding. This is because the ISI must be subtracted immediately, which doesn't allow for any decoding delay. Hence the error propagation seriously degrades performance on channels with low SNR. We can address this problem by introducing some delay in the feedback path to allow for channel decoding [11] or by turbo equalization, described in the next section. A systematic treatment of the DFE with coding can be found in [12, 13]. Moreover, the DFE structure can be generalized to encompass MIMO channels [14].

11.5.3 Turbo and Deep-Learning Equalization

The DFE and Viterbi equalizers are commonly used in practice due to their excellent performance improvements over linear equalizers. In addition, two more recent methods of nonlinear equalization have been developed: turbo equalization, and equalization based on deep learning. The turbo equalizer, introduced in [15], is based on the turbo decoding principle introduced in Section 8.6. This equalizer iterates between a maximum a posteriori (MAP) equalizer and a decoder to determine the transmitted symbol. The MAP equalizer computes the a posteriori probability (APP) of the transmitted symbol given the past channel outputs. The decoder computes the log likelihood ratio (LLR) associated with the transmitted symbol given past channel outputs. The APP and LLR comprise the soft information exchanged between the equalizer and decoder in the turbo iteration. After some number of iterations, the turbo equalizer converges on its estimate of the transmitted symbol. The original turbo equalization principle has been extended to iterate between different types of equalizers and decoding algorithms, as described in [18].

A deep-learning equalizer (DLE) relies on training a deep neural network in the receiver via a massive transmitted data set. Once the training is complete, the trained network is used to estimate the received data sequence. A DLE does not require CSI and, in fact, it doesn't require a linear filter model $H(f)$ for the channel as, in general, it is not attempting to estimate and then remove ISI introduced by $H(f)$ from the received data sequence. Rather, the deep network learns the transmitted data sequence directly based on its past learning. DLEs that do not rely on a model for the channel are called *data-driven* DLEs. These DLEs have demonstrated good performance in mitigating the impact of ISI in several systems where linear models for the channel are elusive or inaccurate, such as in molecular, underwater acoustic, and optical communication systems [16]. The other type of DLE is *model-based*, where some characteristics of the underlying channel are incorporated into the equalizer. For example, a model-based DLE may assume that the channel $H(f)$ has a given delay spread but is otherwise unknown; this structure for the channel is then incorporated into the DLE algorithm. This model-based approach was used in [17] to integrate a deep neural network for learning and tracking channel model parameters assuming a given channel delay spread into the Viterbi algorithm. The method was shown to outperform conventional Viterbi decoding under channel estimation error. In general model-based DLEs have lower complexity and training requirements than data-driven DLEs as they are only being trained on part of the system rather than the entire end-to-end system. In addition, model-based DLEs can take advantage of algorithms that are known to be optimal for specific communication blocks within an end-to-end system, such as Viterbi decoding, while simultaneously benefiting from the learning capabilities of deep networks within communication blocks where the best method for signal processing, particularly under nonideal conditions such as channel estimation error and complexity constraints, are unknown.

11.6 Adaptive Equalizers: Training and Tracking

Most equalizers are designed based on a known value of the combined impulse response $h(t) = g(t) * c(t)$. Since the channel $c(t)$ in generally not known when the receiver is designed, the equalizer must be tunable so it can adjust to different values of $c(t)$. Moreover, since in wireless channels $c(t) = c(\tau, t)$ will change over time, the system must periodically estimate the channel $c(t)$ and update the equalizer coefficients accordingly. This process is called *equalizer training* or *adaptive equalization* [21, 22]. The equalizer can also use the detected data to adjust the equalizer coefficients, a process known as *equalizer tracking*. However, *blind equalizers* do not use training: they learn the channel response via the detected data only [23, 24, 25, 26] or, in the case of DLE-based equalization, they estimate the received sequence without learning the channel response.

During training, the coefficients of the equalizer are updated at time k based on a known training sequence $[d_{k-M}, \ldots, d_k]$ that has been sent over the channel. The length $M + 1$ of the training sequence depends on the number of equalizer coefficients that must be determined and the convergence speed of the training algorithm.

Table 11.1: Equalizer training and tracking characteristics

Algorithm	No. of multiply operations per symbol time T_s	Complexity	Training convergence time	Tracking
LMS	$2N+1$	Low	Slow ($> 10NT_s$)	Poor
MMSE	N^2 to N^3	Very high	Fast ($\sim NT_s$)	Good
RLS	$2.5N^2 + 4.5N$	High	Fast ($\sim NT_s$)	Good
Fast Kalman DFE	$20N + 5$	Fairly low	Fast ($\sim NT_s$)	Good
Square-root LS DFE	$1.5N^2 + 6.5N$	High	Fast ($\sim NT_s$)	Good

Note that the equalizer must be retrained when the channel decorrelates – that is, at least every T_c seconds, where T_c is the channel coherence time. Thus, if the training algorithm is slow relative to the channel coherence time then the channel may change before the equalizer can learn the channel. Specifically, if $(M+1)T_s > T_c$ then the channel will decorrelate before the equalizer has finished training. In this case equalization is not an effective countermeasure for ISI, and some other technique (e.g., multicarrier modulation or CDMA) is needed.

Let $\{\hat{d}_k\}$ denote the bit decisions output from the equalizer given a transmitted training sequence $\{d_k\}$. Our goal is to update the N equalizer coefficients at time $k+1$ based on the training sequence we have received up to time k. We denote these updated coefficients as $\{w_{-L}(k+1), \ldots, w_L(k+1)\}$. We will use the MMSE as our criterion to update these coefficients; that is, we will choose $\{w_{-L}(k+1), \ldots, w_L(k+1)\}$ as the coefficients that minimize the MSE between d_k and $\hat{d}_k$. Recall that $\hat{d}_k = \sum_{i=-L}^{L} w_i(k+1) y_{k-i}$, where $y_k = y[k]$ is the output of the sampler in Figure 11.2 at time k with the known training sequence as input. The $\{w_{-L}(k+1), \ldots, w_L(k+1)\}$ that minimize MSE are obtained via a Weiner filter [4, 6]. Specifically,

$$\mathbf{w}(k+1) = \{w_{-L}(k+1), \ldots, w_L(k+1)\} = \mathbf{R}^{-1}\mathbf{p}, \tag{11.55}$$

where $\mathbf{p} = d_k[y_{k+L} \ldots y_{k-L}]^T$ and

$$\mathbf{R} = \begin{bmatrix} |y_{k+L}|^2 & y_{k+L}y_{k+L-1}^* & \cdots & y_{k+L}y_{k-L}^* \\ y_{k+L-1}y_{k+L}^* & |y_{k+L-1}|^2 & \cdots & y_{k+L-1}y_{k-L}^* \\ \vdots & \ddots & \ddots & \vdots \\ y_{k-L}y_{k+L}^* & \cdots & \cdots & |y_{k-L}|^2 \end{bmatrix}. \tag{11.56}$$

Note that finding the optimal tap updates in this case requires a matrix inversion, which requires N^2 to N^3 multiply operations on each iteration (each symbol time T_s). In exchange for its complexity, the convergence of this algorithm is very fast; it typically converges in around N symbol times for N the number of equalizer tap weights.

If complexity is an issue then the large number of multiply operations needed for MMSE training can be prohibitive. A simpler technique is the least mean square (LMS) algorithm [4, Chap. 11.1.2]. In this algorithm the tap weight vector $\mathbf{w}(k+1)$ is updated linearly as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \triangle\epsilon_k[y_{k+L}^* \ldots y_{k-L}^*], \tag{11.57}$$

where $\epsilon_k = d_k - \hat{d}_k$ is the error between the bit decisions and the training sequence and Δ is the step size of the algorithm. The choice of Δ dictates the convergence speed and stability of the algorithm. For small values of Δ the convergence is very slow; it takes many more than N bits for the algorithm to converge to the proper equalizer coefficients. However, if Δ is chosen to be large then the algorithm can become unstable, basically skipping over the desired tap weights at every iteration. Thus, for good performance of the LMS algorithm, Δ

is typically small and convergence is typically slow. However, the LMS algorithm exhibits significantly reduced complexity compared to the MMSE algorithm, since the tap updates require only about $2N+1$ multiply operations per iteration. Thus, the complexity is linear in the number of tap weights. Other algorithms – such as the root least squares (RLS), square-root least squares, and fast Kalman – provide various trade-offs in terms of complexity and performance that lie between the two extremes of the LMS algorithm (slow convergence but low complexity) and the MMSE algorithm (fast convergence but high complexity). A description of these other algorithms is given in [4, Chap. 11]. Table 11.1 summarizes the specific number of multiply operations and the relative convergence rate of all these algorithms.

Note that the symbol decisions $\hat{d}_k$ output from the equalizer are typically passed through a threshold detector to round the decision to the nearest constellation point. The resulting roundoff error can be used to adjust the equalizer coefficients during data transmission, a process called equalizer tracking. Tracking is based on two premises: (i) that if the roundoff error is nonzero then the equalizer is not perfectly trained; and (ii) that the roundoff error can be used to adjust the channel estimate inherent in the equalizer. The procedure works as follows. The equalizer output bits $\hat{d}_k$ and threshold detector output bits $\hat{\hat{d}}_k$ are used to adjust an estimate of the equivalent lowpass composite channel $F(z)$. In particular, the coefficients of $F(z)$ are adjusted to minimize the MSE between $\hat{d}_k$ and $\hat{\hat{d}}_k$, using the same MMSE procedures described previously. The updated version of $F(z)$ is then taken to equal the composite channel and used to update the equalizer coefficients accordingly. More details can be found in [4, 6]. Table 11.1 also includes a summary of the training and tracking characteristics for the different algorithms as a function of the number of taps N. Note that the fast Kalman and square-root LS may be unstable in their convergence and tracking, which is the price to be paid for their fast convergence with relatively low complexity.

Example 11.5:

Consider a five-tap equalizer that must retrain every $.5T_c$, where T_c is the coherence time of the channel. Assume the transmitted signal is BPSK ($T_s = T_b$) with a rate of 1 Mbps for both data and training sequence transmission. Compare the length of training sequence required for the LMS equalizer versus the fast Kalman DFE. For an 80-Hz Doppler, by how much is the data rate reduced in order to enable periodic training for each of these equalizers? How many operations do these two equalizers require for this training?

Solution: The equalizers must retrain every $.5T_c = .5/B_d = .5/80 = 6.25$ ms. For a data rate of $R_b = 1/T_b = 1$ Mbps, Table 11.1 shows that the LMS algorithm requires $10NT_b = 50 \cdot 10^{-6}$ seconds for training to converge and the fast Kalman DFE requires $NT_b = 50 \cdot 10^{-5}$ seconds. If training occurs every 6.25 ms then the fraction of time the LMS algorithm uses for training is $50 \cdot 10^{-6}/6.25 \cdot 10^{-3} = .008$. Thus, the effective data rate becomes $(1 - .008)R_b = .992$ Mbps. The fraction of time used by the fast Kalman DFE for training is $50 \cdot 10^{-5}/6.25 \cdot 10^{-3} = .0008$, resulting in an effective data rate of $(1 - .0008)R_b = .9992$ Mbps. The LMS algorithm requires approximately $(2N + 1) \cdot 10N = 550$ operations for training per training period, whereas the fast Kalman DFE requires $(20N + 5) \cdot N = 525$ operations. With processor technology today, this is not a significant difference in terms of CPU requirements.

Chapter 11 Problems

1. Design a continuous-time equivalent lowpass equalizer $H_{\text{eq}}(f)$ to completely remove the ISI introduced by a channel with equivalent lowpass impulse response $H(f) = 1/f$. Assume your transmitted signal has a bandwidth of 100 kHz. Assuming a channel with equivalent lowpass AWGN of PSD N_0, find the noise power at the output of your equalizer within the 100-kHz bandwidth of interest. Will this equalizer improve system performance?

2. This problem investigates the interference generated by ISI and the noise enhancement that occurs in zero-forcing equalization. Consider two multipath channels: the first channel has impulse response

$$h_1(t) = \begin{cases} 1 & 0 \le t < T_m, \\ 0 & \text{else;} \end{cases}$$

the second channel has impulse response

$$h_2(t) = e^{-t/T_m},\ 0 \le t < \infty.$$

   (a) Assume that the transmitted signal $s(t)$ is an infinite sequence of impulses with amplitude A and time separation $T_b = T_m/2$; that is, $s(t) = \sum_{n=-\infty}^{\infty} A\delta(t - nT_b)$. Calculate the average ISI power over a bit time T_b.

   (b) Let $T_m = 10\ \mu s$. Suppose a BPSK signal is transmitted over a channel with impulse response $h_1(t)$. Find the maximum data rate that can be sent over the channel with zero ISI under BPSK modulation with rectangular pulse shaping of pulse width $T = 1\ \mu s$. How would this answer change if the signal bandwidth were restricted to 100 kHz?

3. Consider a channel with equivalent lowpass impulse response

$$h(t) = \begin{cases} e^{-t/\tau} & t \ge 0, \\ 0 & \text{else,} \end{cases}$$

where $\tau = 6\ \mu s$. The channel also has equivalent lowpass AWGN with power spectral density N_0.

   (a) What is the frequency response of the equivalent lowpass continuous-time ZF linear equalizer for this channel? Assume no matched filter or pulse shaping.

   (b) Suppose we transmit a 30-kHz signal over this channel, and assume that the frequency response of the signal is a rectangular pulse shape. What is the ratio of SNR with equalization to SNR without equalization in the bandwidth of our transmitted signal? *Hint*: Recall that a stationary random process with power spectral density $P(f)$ has total power $\int P(f)df$ and that, if this process is passed through a filter $G(f)$, then the output process has power spectral density $P(f)|G(f)|^2$.

   (c) Approximate the equivalent lowpass MMSE equalizer for this channel using a discrete-time transversal filter with three taps. Use any method you like as long as it reasonably approximates the time-domain response of the MMSE equalizer.

4. Consider an FIR ZF equalizer with tap weights $w_i = c_i$, where $\{c_i\}$ is the inverse z-transform of $1/F(z)$. Show that this choice of tap weights minimizes

$$\left| \frac{1}{F(z)} - (w_0 + w_1 z^{-1} + \cdots + w_N z^{-N}) \right|^2$$

at $z = e^{j\omega}$.

5. Consider a communication system where the modulated signal $s(t)$ has power 10 mW, carrier frequency f_c, and passband bandwidth $B_s = 40$ MHz. The signal $s(t)$ passes through a frequency-selective fading channel with symmetric frequency response $H(f) = H(-f)$, where

$$H(f) = \begin{cases} 1 & f_c - 20 \text{ MHz} \le f < f_c - 10 \text{ MHz}, \\ .5 & f_c - 10 \text{ MHz} \le f < f_c, \\ 2 & f_c \le f < f_c + 10 \text{ MHz}, \\ .25 & f_c + 10 \text{ MHz} \le f < f_c + 20 \text{ MHz}, \\ 0 & \text{else.} \end{cases}$$

The received signal is $y(t) = s(t) * h(t) + n(t)$, where $n(t)$ is AWGN with PSD $N_0/2 = 10^{-12}$ W/Hz.

   (a) Suppose $y(t)$ is passed through a continuous-time (passband) ZF equalizer. Find the frequency response $H_{\text{eq}}(f)$ for this equalizer within the bandwidth of interest ($f_c \pm 20$ MHz).

   (b) For the equalizer of part (a), find the SNR at the equalizer output.

   (c) Suppose the symbol time for $s(t)$ is $T_s = .5/B_s$ and assume no restrictions on the constellation size. Find the maximum data rate that can be sent over the channel with the ZF equalizer of part (a) such that $P_b < 10^{-3}$.

6. Consider an ISI channel with received signal after transmission through the channel given by

$$y(t) = \sum_{i=-\infty}^{\infty} x_i f(t - iT),$$

where $x_i = \pm 1$ and $f(t)$ is the composite impulse response consisting of the pulse-shaping filter, the equivalent lowpass channel, and the receiver matched filter. Assume that $f(t) = \sin(\pi t/T)/(\pi t/T)$, which satisfies the Nyquist criterion for zero ISI. There are two difficulties with this pulse shape: first, it has a rectangular spectrum, which is difficult to implement in practice. Second, the tails of the pulse decay as $1/t$, so timing error leads to a sequence of ISI samples that do not converge. For parts (b) and (c) of this problem we make the assumption that $f(t) = 0$ for $|t| > NT$, where N is a positive integer. This is not strictly correct, since it would imply that $f(t)$ is both time limited and bandlimited. However, it is a reasonable approximation in practice.

   (a) Show that the folded spectrum of $f(t)$ is flat.

   (b) Suppose that (because of timing error) the signal is sampled at $t = kT + t_0$, where $t_0 < T$. Calculate the response $y_k = y(kT + t_0)$ and separate your answer into the desired term and the ISI terms.

(c) Assume that the polarities of the x_i are such that every term in the ISI is positive (worst-case ISI). Under this assumption, show that the ISI term from part (b) is

$$\text{ISI} \approx \frac{2}{\pi} \sin\left(\frac{\pi t_0}{T}\right) \sum_{n=1}^{N} \frac{n}{n^2 - t_0^2/T^2}$$

and therefore ISI $\to \infty$ as $N \to \infty$.

7. Let $g(t) = \text{sinc}(t/T_s), |t| < T_s$. Find the matched filter $g_m(t)$ for $g(t)$. Find the noise-whitening filter $1/G_m^*(1/z^*)$ for this system that must be used in an MMSE equalizer to whiten the noise.

8. Show that the minimum MSE (11.35) for an IIR MMSE equalizer can be expressed in terms of the folded spectrum $F_\Sigma(f)$ as

$$J_{\min} = T \int_{-.5/T}^{.5/T} \frac{N_0}{F_\Sigma(f) + N_0}\, df.$$

9. Show that the gradient of tap weights associated with the MMSE equalizer is given by

$$\nabla_{\mathbf{w}} J = \left(\frac{\partial J}{\partial w_0}, \ldots, \frac{\partial J}{\partial w_N}\right) = 2\mathbf{w}^T \mathbf{M}_{\mathbf{v}} - 2\mathbf{v}_{\mathbf{d}}.$$

Set this equal to zero and solve for the optimal tap weights to obtain

$$\mathbf{w}_{\text{opt}} = (\mathbf{M}_{\mathbf{v}}^T)^{-1} \mathbf{v}_{\mathbf{d}}^H.$$

10. Show that the MMSE $J_{\min}$ for an IIR MMSE equalizer, given by (11.40), satisfies $0 \le J_{\min} \le 1$.

11. Compare the value of the minimum MSE, $J_{\min}$, under both MMSE equalization and DF equalization for a channel with three-tap discrete-time equivalent model $C(z) = 1 + .5z^{-1} + .3z^{-2}$.

12. This problem illustrates (i) the noise enhancement of zero-forcing equalizers and (ii) how this enhancement can be mitigated using an MMSE approach. Consider a frequency-selective fading channel with equivalent lowpass frequency response

$$H(f) = \begin{cases} 1 & 0 \le |f| < 10 \text{ kHz}, \\ 1/2 & 10\text{kHz} \le |f| < 20 \text{ kHz}, \\ 1/3 & 20\text{kHz} \le |f| < 30 \text{ kHz}, \\ 1/4 & 30\text{kHz} \le |f| < 40 \text{ kHz}, \\ 1/5 & 40\text{kHz} \le |f| < 50 \text{ kHz}, \\ 0 & \text{else}. \end{cases}.$$

The frequency response is symmetric in positive and negative frequencies. Assume an AWGN channel with equivalent lowpass noise PSD $N_0 = 10^{-9}$.

(a) Find a ZF analog equalizer that completely removes the ISI introduced by $H(f)$.

(b) Find the total noise power at the output of the equalizer from part (a).

(c) Assume an MMSE analog equalizer of the form $H_{\text{eq}}(f) = 1/(H(f) + \alpha)$. Find the total noise power at the output of this equalizer for an AWGN input with PSD N_0 for $\alpha = .5$ and for $\alpha = 1$.

(d) Describe qualitatively two effects on a signal that is transmitted over channel $H(f)$ and then passed through the MMSE equalizer $H_{\text{eq}}(f) = 1/(H(f)+\alpha)$ with $\alpha > 0$. What design considerations should go into the choice of α?

(e) What happens to the total noise power for the MMSE equalizer in part (c) as $\alpha \to \infty$? What is the disadvantage of letting $\alpha \to \infty$ in this equalizer design?

(f) For the equalizer design of part (c), suppose the system has a data rate of 100 kbps and that your equalizer requires a training sequence of 1000 bits to train. What is the maximum channel Doppler such that the equalizer coefficients converge before the channel decorrelates?

13. Why does an equalizer that tracks the channel during data transmission still need to train periodically? Specify two benefits of tracking.

14. Assume a four-tap equalizer that must retrain every $.5T_c$, where T_c is the channel coherence time. If a digital signal processing chip can perform 10 million multiplications per second and if the convergence rates for the LMS DFE algorithm and the RLS algorithm are, respectively, 1000 iterations (symbol times) and 50 iterations, then what is the maximum data rate for both equalizers assuming BPSK modulation and Doppler spread $B_D = 100$ Hz? For $B_D = 1000$ Hz? Assume that the transmitting speed is equal for training and information sequences.

15. In this problem we seek the procedure for updating the channel estimate during tracking. Find the formula for updating the channel coefficients corresponding to the channel $H(z)$ based on minimizing the MSE between $\hat{d}_k$ and $\hat{\hat{d}}_k$.

16. Ultrawideband (UWB) systems spread a data signal and its corresponding power over a very wide bandwidth so that the power per hertz of the signal is small (typically below the noise floor). Hence such systems coexist with other systems without causing them much interference. Consider a baseband UWB system with BPSK modulation. The data bits are modulated with a rectangular pulse $g(t)$, where $g(t)$ has a very narrow time duration T as compared to the bit time T_b. For this problem we assume $T = 10^{-9}$. Thus, a UWB signal with BPSK modulation would have the form $s(t) = \sum_n d_n g(t - nT_b)$, where d_n takes the value ± 1 and $T_b \gg T$ is the bit time. A sketch of $s(t)$ with a data sequence of alternating 1-bits and 0-bits is shown in Figure 11.5.

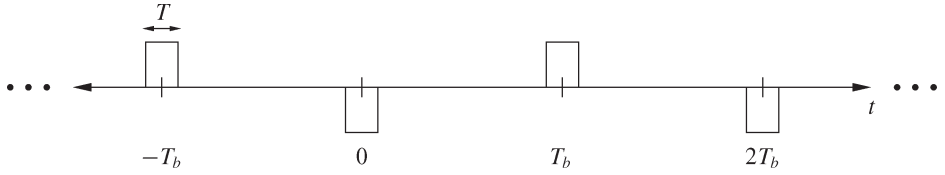


Figure 11.5: UWB signal form for Problem 11-16.

(a) What is the approximate bandwidth of $s(t)$ if $T_b = 10^{-5}$?

(b) One of the selling points of UWB is that its signals do not experience flat fading in typical channels. Consider a single-bit transmission $s(t) = d_0 g(t)$. Suppose $s(t)$ is transmitted through a channel that follows a two-ray model $h(t) = \alpha_0\delta(t)+\alpha_1\delta(t-\tau)$. Sketch the channel output for $\tau \ll T$ and $\tau \gg T$. Which of your two sketches is more likely to depict the output of a real wireless channel? Why does this imply that UWB signals don't typically experience flat fading?

(c) Consider a channel with a multipath delay spread of $T_m = 20\ \mu s$. For Figure 11.5, what is the *exact* maximum data rate that can be sent over this channel with no ISI? Is the bandwidth of $s(t)$ in the figure less than the channel coherence bandwidth at this data rate?

(d) Let $F(z) = \alpha_0 + \alpha_1 z^{-1} + \alpha_2 z^{-2}$ denote the composite impulse response of the transmitter, channel, and matched filter in this UWB system. Find a two-tap digital equalizer $H_{\text{eq}}(z) = w_0 + w_1 z^{-1}$ that approximates an IIR zero-forcing equalizer for $F(z)$. Any reasonable approximation is fine as long as you can justify it.

(e) For the equalizer designed in part (d), suppose the system has a data rate of 100 kbps and that your equalizer requires a training sequence of 1000 bits to train. What is the maximum channel Doppler such that the equalizer coefficients converge before the channel decorrelates?

Bibliography

[1] E. H. Satorius and S. T. Alexander, "Channel equalization using adaptive lattice algorithms," *IEEE Trans. Commun*., pp. 899–905, June 1979.

[2] F. Ling and J. Proakis, "Adaptive lattice decision feedback equalizers – Their performance and application to time-variant multipath channels," *IEEE Trans. Commun*., pp. 348–56, April 1985.

[3] J. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Signal Proc*., pp. 304–37, April 1984.

[4] J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001.

[5] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems", *IEEE Communications Magazine*, vol. 40, no. 4, pp. 58-66, April 2002

[6] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[7] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of inter-symbol interference," *IEEE Trans. Inform. Theory*, pp. 363–78, May 1972.

[8] B. Sklar, "How I learned to love the trellis," *IEEE Signal Proc. Mag*., pp. 87–102, May 2003.

[9] R. Price, "Nonlinearly feedback-equalized PAM vs. capacity," *Proc. IEEE Internat. Conf. Comun*., pp. 22.12–22.17, June 1972.

[10] J. Salz, "Optimum mean-square decision feedback equalization," *Bell System Tech. J*., pp. 1341–73, October 1973.

[11] M. V. Eyuboglu, "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interleaving," *IEEE Trans. Commun*., pp. 401–9, April 1988.

[12] J. M. Cioffi, G. P. Dudevoir, V. Eyuboglu, and G. D. Forney, Jr., "MMSE decision-feedback equalizers and coding. Part I: Equalization results," *IEEE Trans. Commun*., pp. 2582–94, October 1995.

[13] J. M. Cioffi, G. P. Dudevoir, V. Eyuboglu, and G. D. Forney, Jr., "MMSE decision-feedback equalizers and coding. Part II: Coding results," *IEEE Trans. Commun*., pp. 2595–2604, October 1995.

[14] J. M. Cioffi and G. D. Forney, Jr., "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in A. Paulraj, V. Roychowdhury, and C. Schaper (Eds.), *Communication, Computation, Control, and Signal Processing*, pp. 79–127, Kluwer, Boston, 1997.

[15] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *Euro. Trans. Telecommun.*, pp. 507–11, September/October 1995.

[16] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5663-5678, Nov. 2018.

[17] N. Shlezinger, N. Farsad, Y. C. Eldar and A. J. Goldsmith, "ViterbiNet: A deep learning based Viterbi algorithm for symbol detection," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3319-3331, May 2020.

[18] M. Tüchler and A. C. Singer, "Turbo equalization: An overview," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 920-952, Feb. 2011.

[19] M. Tomlinson, "A new automatic equalizer employing modulo arithmetic," *Elec. Lett.*, 7, pp. 138–9, 1971.

[20] H. Harashima and H. Miyakawa, "Matched-transmission techniques for channels with intersymbol interference," *IEEE Trans. Commun.*, pp. 774–80, August 1972.

[21] J. G. Proakis, "Adaptive equalization for TDMA digital mobile radio," *IEEE Trans. Veh. Tech.*, pp. 333–41, May 1991.

[22] S. U. Qureshi, "Adaptive equalization," *Proc. IEEE*, pp. 1349–87, September 1985.

[23] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, pp. 871–83, August 1984.

[24] C. R. Johnson Jr., "Admissibility in blind adaptive channel equalization," *IEEE Control Syst. Mag.*, pp. 3–15, January 1991.

[25] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, pp. 1927–50, October 1998.

[26] L. Tong, G. Zu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, pp. 340–9, March 1994.

Chapter 12

Multicarrier Modulation

The basic idea of multicarrier modulation is to divide a wideband channel into multiple narrowband subchannels. Specifically, the transmitted bitstream is divided into many different substreams, with each substream transmitted over a different subchannel. Typically the subchannels are orthogonal under ideal propagation conditions. The data rate on each of the subchannels is much less than the total data rate, and the corresponding subchannel bandwidth is much less than the total system bandwidth. The number of substreams is chosen to ensure that each subchannel has a bandwidth less than the coherence bandwidth of the channel, so the subchannels experience relatively flat fading. Thus, the intersymbol interference on each subchannel is small. The subchannels in multicarrier modulation need not be contiguous, so a large continuous block of spectrum is not needed for high-rate multicarrier communications. Moreover, multicarrier modulation is efficiently implemented digitally. In this discrete implementation, called orthogonal frequency division multiplexing (OFDM), the ISI can be completely eliminated through the use of a cyclic prefix.

Multicarrier modulation is currently used in many wireless systems. However, it is not a new technique: it was first used for military HF radios in the late 1950s and early 1960s. Starting around 1990 [1], multicarrier modulation has been used in many diverse wired and wireless applications, including digital audio and video broadcasting in Europe [2], digital subscriber lines (DSL) using discrete multitone [3, 4, 5], and the most recent generation of wireless LANs [5, 6, 7]. There are also a number of newly emerging uses for multicarrier techniques, including fixed wireless broadband services [8, 9], mobile wireless broadband known as FLASH-OFDM [10], and even for ultrawideband radios, where multiband OFDM is one of two competing proposals for the IEEE 802.15 ultrawideband standard. Multicarrier modulation is also a candidate for the air interface in next-generation cellular systems [11, 12].

The multicarrier technique can be implemented in multiple ways, including vector coding [13, 14] and OFDM [13, 15], all of which are discussed in this chapter. These techniques have subtle differences, but all are based on the same premise of breaking a wideband channel into multiple parallel narrowband channels by means of an orthogonal channel partition.

There is some debate as to whether multicarrier or single-carrier modulation is better for ISI channels with delay spreads on the order of the symbol time. It is claimed in [2] that, for some mobile radio applications, single-carrier modulation with equalization has roughly the same performance as multicarrier modulation with channel coding, frequency-domain interleaving, and weighted maximum likelihood decoding. Adaptive loading was not taken into account in [2], though it has the potential to significantly improve multicarrier performance [16]. But there are other problems with multicarrier modulation that impair its performance, most significantly frequency offset and timing jitter, which degrade the orthogonality of the subchannels. In addition, the peak-to-average power ratio of multicarrier systems is significantly higher than that of single-carrier systems, which is a serious problem when nonlinear amplifiers are used. Trade-offs between multicarrier and single-carrier block transmission

systems with respect to these impairments are discussed in [17].

Despite these challenges, multicarrier techniques are common in high–data-rate wireless systems with moderate to large delay spread, as they have significant advantages over time-domain equalization. In particular, the number of taps required for an equalizer with good performance in a high–data-rate system is typically large. Thus, these equalizers are quite complex. Moreover, it is difficult to maintain accurate weights for a large number of equalizer taps in a rapidly varying channel. For these reasons, most emerging high-rate wireless systems use either multicarrier modulation or spread spectrum instead of equalization to compensate for ISI.

12.1 Data Transmission Using Multiple Carriers

The simplest form multicarrier modulation divides the data stream into multiple substreams to be transmitted over different orthogonal subchannels centered at different subcarrier frequencies. The number of substreams is chosen to make the symbol time on each substream much greater than the delay spread of the channel or, equivalently, to make the substream bandwidth less than the channel coherence bandwidth. This ensures that the substreams will not experience significant ISI.

Consider a linearly modulated system with data rate R and bandwidth B. The coherence bandwidth for the channel is assumed to be $B_c < B$, so the signal experiences frequency-selective fading. The basic premise of multicarrier modulation is to break this wideband system into N linearly modulated subsystems in parallel, each with subchannel bandwidth $B_N = B/N$ and data rate $R_N \approx R/N$. For N sufficiently large, the subchannel bandwidth $B_N = B/N \ll B_c$, which ensures relatively flat fading on each subchannel. This can also be seen in the time domain: the symbol time T_N of the modulated signal in each subchannel is proportional to the reciprocal of the subchannel bandwidth $1/B_N$. So $B_N \ll B_c$ implies that $T_N \approx 1/B_N \gg 1/B_c \approx T_m$, where T_m denotes the delay spread of the channel. Thus, if N is sufficiently large, the symbol time is much greater than the delay spread, so each subchannel experiences little ISI degradation.

Figure 12.1 illustrates a multicarrier transmitter.[1] The bit stream is divided into N substreams via a serial-to-parallel converter. The nth substream is linearly modulated (typically via QAM or PSK) relative to the subcarrier frequency f_n and occupies bandwidth B_N. We assume coherent demodulation of the subcarriers so the subcarrier phase is neglected in our analysis. If we assume raised cosine pulses for $g(t)$ we get a symbol time $T_N = (1+\beta)/B_N$ for each substream, where β is the rolloff factor of the pulse shape. The modulated signals associated with all the subchannels are summed together to form the transmitted signal, given as

$$s(t) = \sum_{i=0}^{N-1} s_i \, g(t) \cos(2\pi f_i t + \phi_i), \tag{12.1}$$

where s_i is the complex symbol associated with the ith subcarrier and ϕ_i is the phase offset of the ith carrier. For nonoverlapping subchannels we set $f_i = f_0 + i(B_N)$, $i = 0, \ldots, N-1$. The substreams then occupy orthogonal subchannels with bandwidth B_N, yielding a total bandwidth $NB_N = B$ and data rate $NR_N \approx R$. Thus, this form of multicarrier modulation does not change the data rate or signal bandwidth relative to the original system, but it almost completely eliminates ISI for $B_N \ll B_c$.

The receiver for this multicarrier modulation is shown in Figure 12.2. Each substream is passed through a narrowband filter (to remove the other substreams), demodulated, and combined via a parallel-to-serial converter to form the original data stream. Note that the ith subchannel will be affected by flat fading corresponding to a channel gain $\alpha_i = H(f_i)$.

[1] In practice, the complex symbol s_i would have its real part transmitted over the in-phase signaling branch and its imaginary part transmitted over the quadrature signaling branch. For simplicity we illustrate multicarrier transmission based on sending a complex symbol along the in-phase signaling branch.

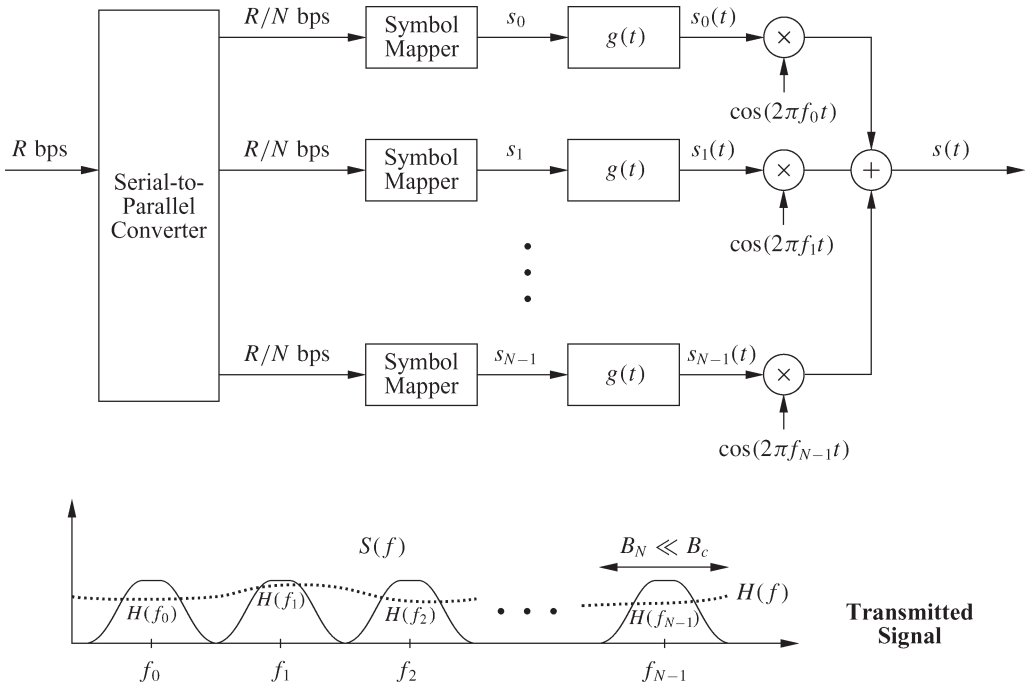


Figure 12.1: Multicarrier transmitter.

Although this simple type of multicarrier modulation is easy to understand, it has several significant shortcomings. First, in a realistic implementation, subchannels will occupy a larger bandwidth than under ideal raised cosine pulse shaping because the pulse shape must be time limited. Let ε/T_N denote the additional bandwidth required due to time limiting of these pulse shapes. The subchannels must then be separated by $(1+\beta+\varepsilon)/T_N$, and since the multicarrier system has N subchannels, the bandwidth penalty for time limiting is $\varepsilon N/T_N$. In particular, the total required bandwidth for nonoverlapping subchannels is

$$B = \frac{N(1+\beta+\varepsilon)}{T_N}. \tag{12.2}$$

Thus, this form of multicarrier modulation can be spectrally inefficient. Additionally, near-ideal (and hence expensive) lowpass filters will be required to maintain the orthogonality of the subcarriers at the receiver. Perhaps most importantly, this scheme requires N independent modulators and demodulators, which entails significant expense, size, and power consumption. The next section presents a modulation method that allows subcarriers to overlap and removes the need for tight filtering. Section 12.4 presents the discrete implementation of multicarrier modulation, which eliminates the need for multiple modulators and demodulators.

Example 12.1: Consider a multicarrier system with a total passband bandwidth of 1 MHz. Suppose the system operates in a city with channel delay spread $T_m = 20$ μs. How many subchannels are needed to obtain approximately flat fading in each subchannel?

Solution: The channel coherence bandwidth is $B_c = 1/T_m = 1/.00002 = 50\,\text{kHz}$. To ensure flat fading on each subchannel, we take $B_N = B/N = .1B_c \ll B_c$. Thus, $N = B/.1B_c = 1000000/5000 = 200$ subchannels are needed to ensure flat fading on each subchannel. In discrete implementations of multicarrier modulation, N

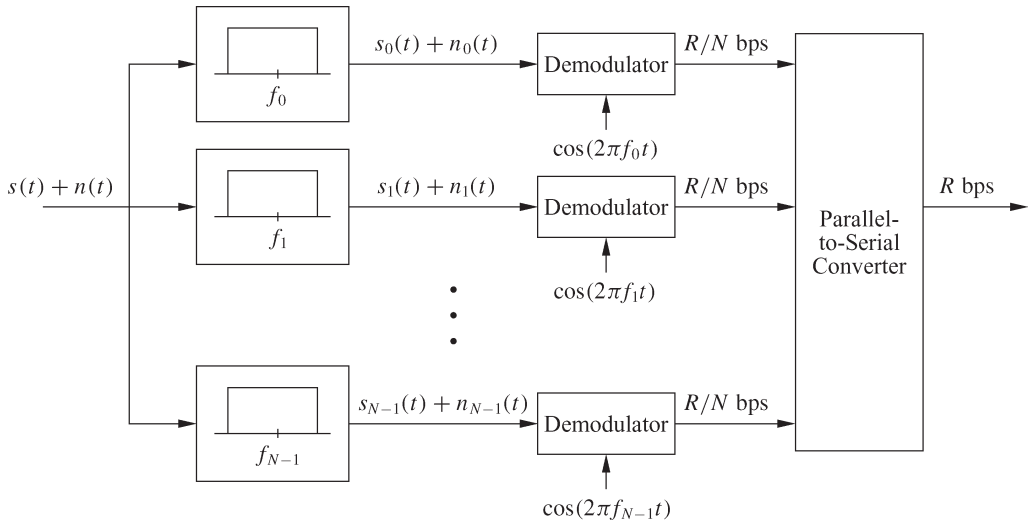


Figure 12.2: Multicarrier receiver.

must be a power of 2 for the DFT (discrete Fourier transform) and IDFT (inverse DFT) operations, in which case $N = 256$ for this set of parameters.

Example 12.2: Consider a multicarrier system with $T_N = .2$ ms: $T_N \gg T_m$ for T_m the channel delay spread, so each subchannel experiences minimal ISI. Assume the system has $N = 128$ subchannels. If time-limited raised cosine pulses with $\beta = 1$ are used and if the additional bandwidth required (because of time limiting) to ensure minimal power outside the signal bandwidth is $\varepsilon = .1$, then what is the total bandwidth of the system?
Solution: From (12.2),

$$B = \frac{N(1+\beta+\varepsilon)}{T_N} = \frac{128(1+1+.1)}{.0002} = 1.344\,\text{MHz}. \tag{12.3}$$

We will see in the next section that the bandwidth requirements for this system can be substantially reduced by overlapping subchannels.

12.2 Multicarrier Modulation with Overlapping Subchannels

We can improve on the spectral efficiency of multicarrier modulation by overlapping the subchannels. The subcarriers must still be orthogonal so that they can be separated out by the demodulator in the receiver. The subcarriers $\{\cos(2\pi(f_0 + i/T_N)t + \phi_i),\ i = 0, 1, 2, \ldots\}$ form a set of (approximately) orthogonal basis functions on the interval $[0,\ T_N]$ for any set of subcarrier phase offsets $\{\phi_i\}$, since

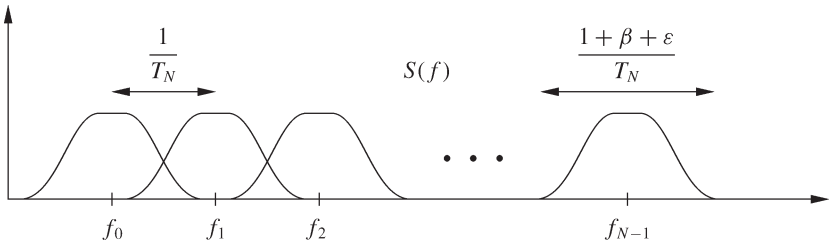


Figure 12.3: Multicarrier with overlapping subcarriers.

$$\begin{aligned}
&\frac{1}{T_N}\int_0^{T_N} \cos\left(2\pi\left(f_0+\frac{i}{T_N}\right)t+\phi_i\right)\cos\left(2\pi\left(f_0+\frac{j}{T_N}\right)t+\phi_j\right)dt \\
&\quad=\frac{1}{T_N}\int_0^{T_N} .5\cos\left(2\pi\frac{(i-j)t}{T_N}+\phi_i-\phi_j\right)dt \\
&\qquad+\frac{1}{T_N}\int_0^{T_N} .5\cos\left(2\pi\left(2f_0+\frac{i+j}{T_N}\right)t+\phi_i+\phi_j\right)dt \\
&\quad\approx\frac{1}{T_N}\int_0^{T_N} .5\cos\left(2\pi\frac{(i-j)t}{T_N}+\phi_i-\phi_j\right)dt \\
&\quad=.5\delta(i-j), \qquad (12.4)
\end{aligned}$$

where the approximation follows because the integral in the third line of (12.4) is approximately zero for $f_0T_N \gg 1$. Moreover, it is easily shown that no set of subcarriers with a smaller frequency separation forms an orthogonal set on $[0, T_N]$ for arbitrary subcarrier phase offsets. This implies that the minimum frequency separation required for subcarriers to remain orthogonal over the symbol interval $[0, T_N]$ is $1/T_N$. Since the carriers are orthogonal, by Section 5.1 the set of functions $\{g(t)\cos(2\pi(f_0 + i/T_N)t + \phi_i),\ i = 0, 1, \ldots, N-1\}$ also form a set of (approximately) orthonormal basis functions for appropriately chosen baseband pulse shapes $g(t)$: the family of raised cosine pulses are a common choice for this pulse shape. Given this orthonormal basis set, even if the subchannels overlap, the modulated signals transmitted in each subchannel can be separated out in the receiver, as we now show.

Consider a multicarrier system where each subchannel is modulated using raised cosine pulse shapes with rolloff factor β. The bandwidth of each subchannel is then $B_N = (1+\beta)/T_N$. The ith subcarrier frequency is set to $(f_0 + i/T_N)$, $i = 0, 1, \ldots, N-1$, for some f_0, so the subcarriers are separated by $1/T_N$. However, the bandwidth of each subchannel is $B_N = (1+\beta)/T_N > 1/T_N$ for $\beta > 0$, so the subchannels overlap. Excess bandwidth due to time windowing will increase the subcarrier bandwidth by an additional ε/T_N. However, β and ε do not affect the total system bandwidth resulting from the subchannel overlap except in the first and last subchannels, as illustrated in Figure 12.3. The total system bandwidth with overlapping subchannels is given by

$$B = \frac{N+\beta+\varepsilon}{T_N} \approx \frac{N}{T_N}, \qquad (12.5)$$

where the approximation holds for N large. Thus, with N large, the impact of β and ε on the total system bandwidth is negligible, in contrast to the required bandwidth of $B = N(1+\beta+\varepsilon)/T_N$ when the subchannels do not overlap.

Example 12.3: Compare the required bandwidth of a multicarrier system with overlapping subchannels versus nonoverlapping subchannels using the same parameters as in Example 12.2.

Solution: In the prior example $T_N = .2$ ms, $N = 128,\ \beta = 1$, and $\varepsilon = .1$. With overlapping subchannels, from (12.5) we have

$$B = \frac{N + \beta + \varepsilon}{T_N} = \frac{128 + 1 + .1}{.0002} = 645.5 \text{ kHz} \approx N/T_N = 640 \text{ kHz}. \tag{12.6}$$

By comparison, in the prior example the required bandwidth with nonoverlapping subchannels was shown to be 1.344 MHz, more than double the required bandwidth when the subchannels overlap.

Clearly, in order to separate out overlapping subcarriers, a different receiver structure is needed than the one shown in Figure 12.2. In particular, overlapping subchannels are demodulated with the receiver structure shown in Figure 12.4, which demodulates the appropriate symbol without interference from overlapping subchannels. Specifically, if the effect of the channel $h(t)$ and noise $n(t)$ are neglected then, for received signal $s(t)$ given by (12.1), the input to each symbol demapper in Figure 12.4 is

$$\begin{aligned}
\hat{s}_i &= \int_0^{T_N} \left(\sum_{j=0}^{N-1} s_j g(t) \cos(2\pi f_j t + \phi_j) \right) g(t) \cos(2\pi f_i t + \phi_i)\, dt \\
&= \sum_{j=0}^{N-1} s_j \int_0^{T_N} g^2(t) \cos\left(2\pi \left(f_0 + \frac{j}{T_N}\right) t + \phi_j\right) \cos\left(2\pi \left(f_0 + \frac{i}{T_N}\right) t + \phi_i\right) dt \\
&= \sum_{j=0}^{N-1} s_j \delta(j - i) \qquad (12.7) \\
&= s_i, \qquad (12.8)
\end{aligned}$$

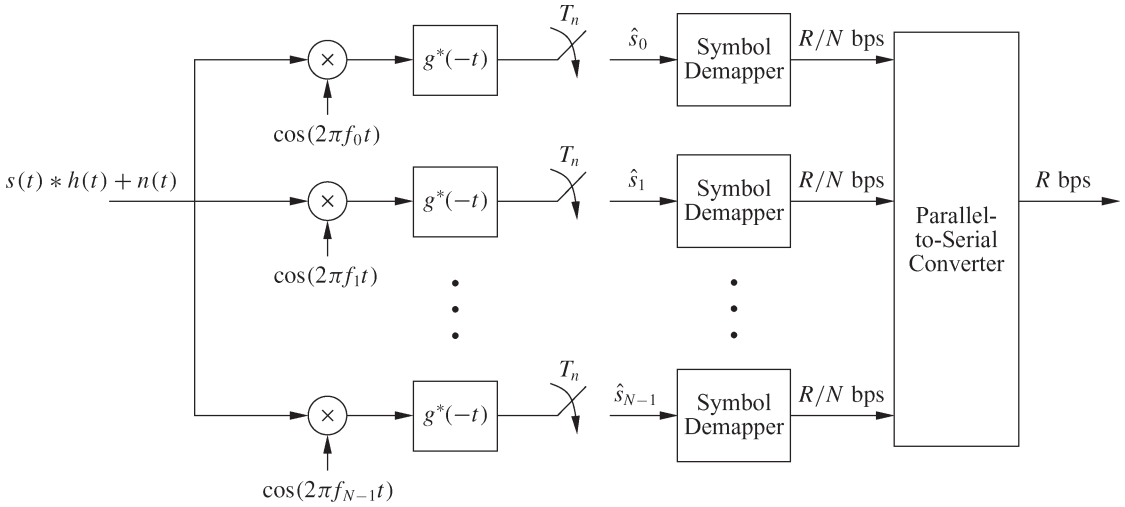


Figure 12.4: Multicarrier receiver for overlapping subcarriers.

where (12.7) follows from the fact that the functions $\{g(t)\cos(2\pi f_j t + \phi_j)\}$ form a set of orthonormal basis functions on $[0, T_N]$. If the channel and noise effects are included, then the symbol in the ith subchannel is scaled

by the channel gain $\alpha_i = H(f_i)$ and corrupted by the noise sample, so $\hat{s}_i = \alpha_i s_i + n_i$, where n_i is additive white Gaussian noise with power $N_0 B_N$. This multicarrier system makes much more efficient use of bandwidth than in systems with nonoverlapping subcarriers. However, since the subcarriers overlap, their orthogonality is compromised by timing and frequency offset. Even when relatively small, these effects can significantly degrade performance, as they cause subchannels to interfere with each other. These effects are discussed in more detail in Section 12.5.2.

12.3 Mitigation of Subcarrier Fading

The advantage of multicarrier modulation is that each subchannel is relatively narrowband, which mitigates the effect of delay spread. However, each subchannel experiences flat fading, which can cause large bit error rates on some of the subchannels. In particular, if the transmit power on subcarrier i is P_i and if the fading on that subcarrier is α_i, then the received signal-to-noise power ratio is $\gamma_i = \alpha_i^2 P_i/(N_0 B_N)$, where B_N is the bandwidth of each subchannel. If α_i is small then the received SNR on the ith subchannel is low, which can lead to a high BER on that subchannel. Moreover, in wireless channels α_i will vary over time according to a given fading distribution, resulting in the same performance degradation as is associated with flat fading for single-carrier systems discussed in Chapter 6. Because flat fading can seriously degrade performance in each subchannel, it is important to compensate for flat fading in the subchannels. There are several techniques for doing this, including coding with interleaving over time and frequency, frequency equalization, precoding, and adaptive loading, all described in subsequent sections. Coding with interleaving is the most common, and it has been adopted as part of the European standards for digital audio and video broadcasting [2, 18]. Moreover, in rapidly changing channels it is difficult to estimate the channel at the receiver and feed this information back to the transmitter. Without channel information at the transmitter, precoding and adaptive loading cannot be done, so only coding with interleaving is effective at fading mitigation.

12.3.1 Coding with Interleaving over Time and Frequency

The basic idea in coding with interleaving over time and frequency is to encode data bits into codewords, interleave the resulting coded bits over both time and frequency, and then transmit the coded bits over different subchannels such that the coded bits within a given codeword all experience independent fading [19]. If most of the subchannels have a high SNR, the codeword will have most coded bits received correctly, and the errors associated with the few bad subchannels can be corrected. Coding across subchannels basically exploits the frequency diversity inherent in a multicarrier system to correct for errors. This technique works well only if there is sufficient frequency diversity across the total system bandwidth. If the coherence bandwidth of the channel is large, then the fading across subchannels will be highly correlated, which will significantly reduce the benefits of coding. Most coding schemes assume channel information in the decoder. Channel estimates are typically obtained by a two-dimensional pilot symbol transmission over both time and frequency [20].

Note that coding with frequency/time interleaving takes advantage of the fact that the data on all the subcarriers are associated with the same user and can therefore be jointly processed. The other techniques for fading mitigation discussed in subsequent sections are all basically flat fading compensation techniques, which apply equally to multicarrier systems as well as to narrowband flat fading single-carrier systems.

12.3.2 Frequency Equalization

In frequency equalization the flat fading α_i on the ith subchannel is basically inverted in the receiver [2]. Specifically, the received signal is multiplied by $1/\alpha_i$, which gives a resultant signal power $\alpha_i^2 P_i/\alpha_i^2 = P_i$. While this removes the impact of flat fading on the signal, it enhances the noise power. Hence the incoming noise signal

is also multiplied by $1/\alpha_i$, so the noise power becomes $N_0 B_N/\alpha_i^2$ and the resultant SNR on the ith subchannel after frequency equalization is the same as before equalization. Therefore, frequency equalization does not really change the performance degradation associated with subcarrier flat fading.

12.3.3 Precoding

Precoding uses the same idea as frequency equalization, except that the fading is inverted at the transmitter instead of at the receiver [21]. This technique requires the transmitter to have knowledge of the subchannel flat fading gains α_i, $i = 0, \ldots, N-1$, which must be obtained through estimation [27]. In this case, if the desired received signal power in the ith subchannel is P_i and if the channel introduces a flat fading gain α_i in the ith subchannel, then under precoding the power transmitted in the ith subchannel is P_i/α_i^2. The subchannel signal is corrupted by flat fading with gain α_i, so the received signal power is $P_i\alpha_i^2/\alpha_i^2 = P_i$, as desired. Note that the channel inversion takes place at the transmitter instead of the receiver, so the noise power remains $N_0 B_N$. Precoding is quite common on wireline multicarrier systems like high–bit-rate digital subscriber lines (HDSL). There are two main problems with precoding in a wireless setting. First, precoding is basically channel inversion, and we know from Section 6.3.5 that inversion is not power efficient in fading channels. In particular, an infinite amount of power is needed for channel inversion on a Rayleigh fading channel. The other problem with precoding is the need for accurate channel estimates at the transmitter, which are difficult to obtain in a rapidly fading channel.

12.3.4 Pulse Shaping

Time-frequency pulse shaping of the transmitted signal can be used in multicarrier modulation to minimize dispersion after transmission through the channel. Various criteria for pulse shape optimization have been considered including suppressing both ISI and ICI [22], maximizing signal-to-interference-plus-noise ratio (SINR) [23], or optimizing spectral efficiency through the use of non-rectangular lattices in the time-frequency domain [24]. An overview of pulse-shape design for OFDM and the resulting performance benefits can be found in [25, 26].

12.3.5 Adaptive Loading

Adaptive loading is based on the adaptive modulation techniques discussed in Chapter 9. It is commonly used on slowly changing channels like digital subscriber lines [16], where channel estimates at the transmitter can be obtained fairly easily. The basic idea is to vary the data rate and power assigned to each subchannel relative to that subchannel gain. As in the case of precoding, this requires knowledge of the subchannel fading $\{\alpha_i,\ i = 0, \ldots, N-1\}$ at the transmitter. In adaptive loading, power and rate on each subchannel are adapted to maximize the total rate of the system using adaptive modulation such as variable-rate variable-power MQAM.

Before investigating adaptive modulation, let us consider the capacity of the multicarrier system with N independent subchannels of bandwidth B_N and subchannel gain $\{\alpha_i,\ i = 0, \ldots, N-1\}$. Assuming a total power constraint P, this capacity is given by[2]

$$C = \max_{P_i: \sum P_i = P} \sum_{i=0}^{N-1} B_N \log_2 \left(1 + \frac{\alpha_i^2 P_i}{N_0 B_N}\right). \tag{12.9}$$

[2]As discussed in Section 4.3.1, this summation is the exact capacity when the α_i are independent. However, in order for the α_i to be independent, the subchannels must be separated by the coherence bandwidth of the channel, which would imply that the subchannels are no longer flat fading. Since the subchannels are designed to be flat fading, the subchannel gains $\{\alpha_i,\ i = 1, \ldots, N\}$ will be correlated, in which case the capacity obtained by summing over the capacity in each subchannel is an upper bound on the true capacity. We will take this bound to be the actual capacity, since in practice the bound is quite tight.

The power allocation P_i that maximizes this expression is a water-filling over frequency given by equation (4.32):

$$\frac{P_i}{P} = \begin{cases} 1/\gamma_c - 1/\gamma_i & \gamma_i \geq \gamma_c, \\ 0 & \gamma_i < \gamma_c, \end{cases} \tag{12.10}$$

for some cutoff value γ_c, where $\gamma_i = \alpha_i^2 P/(N_0 B_N)$. The cutoff value is obtained by substituting the power adaptation formula into the power constraint. The capacity then becomes

$$C = \sum_{i:\gamma_i \geq \gamma_c} B_N \log_2 \left(\frac{\gamma_i}{\gamma_c} \right). \tag{12.11}$$

If we now apply the variable-rate variable-power MQAM modulation scheme described in Section 9.3.2 to the subchannels, then the total data rate is given by

$$R = \sum_{i=0}^{N-1} B_N \log_2 \left(1 + \frac{K\gamma_i P_i}{P} \right), \tag{12.12}$$

where $K = -1.5/\ln(5P_b)$ for P_b the desired target BER in each subchannel. Optimizing this expression relative to the P_i yields the optimal power allocation,

$$\frac{KP_i}{P} = \begin{cases} 1/\gamma_K - 1/\gamma_i & \gamma_i \geq \gamma_K, \\ 0 & \gamma_i < \gamma_K, \end{cases} \tag{12.13}$$

and corresponding data rate,

$$R = \sum_{i:\gamma_i \geq \gamma_K} B_N \log \left(\frac{\gamma_i}{\gamma_K} \right), \tag{12.14}$$

where γ_K is a cutoff fade depth dictated by the power constraint P and K.

12.4 Discrete Implementation of Multicarrier Modulation

Although multicarrier modulation was invented in the 1950s, its requirement for separate modulators and demodulators on each subchannel was far too complex for most system implementations at the time. However, the development of simple and cheap implementations of the discrete Fourier transform and the inverse DFT twenty years later – combined with the realization that multicarrier modulation could be implemented with these algorithms – ignited its widespread use. In this section, after first reviewing the basic properties of the DFT, we illustrate OFDM, which implements multicarrier modulation using the DFT and IDFT.

12.4.1 The DFT and Its Properties

Let $x[n],\ 0 \leq n \leq N-1$, denote a discrete time sequence. The N-point DFT of $x[n]$ is defined [28] as

$$\mathrm{DFT}\{x[n]\} = X[i] \triangleq \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j2\pi ni/N}, \quad 0 \leq i \leq N-1. \tag{12.15}$$

The DFT is the discrete-time equivalent to the continuous-time Fourier transform, because $X[i]$ characterizes the frequency content of the time samples $x[n]$ associated with the original signal $x(t)$. Both the continuous-time

Fourier transform and the DFT are based on the fact that complex exponentials are eigenfunctions for any linear system. The sequence $x[n]$ can be recovered from its DFT using the IDFT:

$$\text{IDFT}\{X[i]\} = x[n] \stackrel{\triangle}{=} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X[i] e^{j2\pi ni/N}, \quad 0 \le n \le N-1. \tag{12.16}$$

The DFT and its inverse are typically performed via hardware using the fast Fourier transform (FFT) and inverse FFT (IFFT).

When an input data stream $x[n]$ is sent through a linear time-invariant discrete-time channel $h[n]$, the output $y[n]$ is the discrete-time convolution of the input and the channel impulse response:

$$y[n] = h[n] * x[n] = x[n] * h[n] = \sum_k h[k]x[n-k]. \tag{12.17}$$

The N-point *circular convolution* of $x[n]$ and $h[n]$ is defined as

$$y[n] = x[n] \circledast h[n] = h[n] \circledast x[n] = \sum_k h[k]x[n-k]_N, \tag{12.18}$$

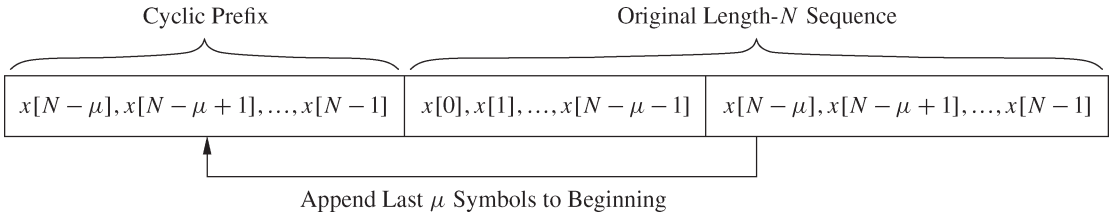


Figure 12.5: Cyclic prefix of length μ.

where $[n-k]_N$ denotes $[n-k]$ modulo N. In other words, $x[n-k]_N$ is a periodic version of $x[n-k]$ with period N. It is easily verified that $y[n]$ given by (12.18) is also periodic with period N. From the definition of the DFT, circular convolution in time leads to multiplication in frequency:

$$\text{DFT}\{y[n] = x[n] \circledast h[n]\} = X[i]H[i], \quad 0 \le i \le N-1, \tag{12.19}$$

where $H[i]$ is the N-point of $\{h[n]\}$. Note that if the sequence $\{h[n]\}$ is of length $k < N$ then it is padded with $N-k$ zeros to obtain a length N for the N-point DFT. By (12.19), if the channel and input are circularly convoluted then, as long as $h[n]$ is known at the receiver, the original data sequence $x[n]$ can be recovered by taking the IDFT of $Y[i]/H[i], 0 \le i \le N-1$. Unfortunately, the channel output is not a circular convolution but a linear convolution. However, the linear convolution between the channel input and impulse response can be turned into a circular convolution by adding a special prefix to the input called a *cyclic prefix*, described in the next section.

12.4.2 The Cyclic Prefix

Consider a channel input sequence $x[n] = x[0], \ldots, x[N-1]$ of length N and a discrete-time channel with finite impulse response (FIR) $h[n] = h[0], \ldots, h[\mu]$ of length $\mu + 1 = T_m/T_s$, where T_m is the channel delay spread and T_s the sampling time associated with the discrete time sequence. The cyclic prefix for $x[n]$ is defined as $\{x[N-\mu], \ldots, x[N-1]\}$: it consists of the last μ values of the $x[n]$ sequence. For each input sequence of length N, these last μ samples are appended to the beginning of the sequence. This yields a new sequence

$\tilde{x}[n], -\mu \leq n \leq N-1$, of length $N+\mu$, where $\tilde{x}[-\mu], \ldots, \tilde{x}[N-1] = x[N-\mu], \ldots, x[N-1], x[0], \ldots, x[N-1]$, as shown in Figure 12.5. Note that with this definition, $\tilde{x}[n] = x[n]_N$ for $-\mu \leq n \leq N-1$, which implies that $\tilde{x}[n-k] = x[n-k]_N$ for $-\mu \leq n-k \leq N-1$.

Suppose $\tilde{x}[n]$ is input to a discrete-time channel with impulse response $h[n]$. The channel output $y[n]$, $0 \leq n \leq N-1$, is then

$$\begin{aligned} y[n] &= \tilde{x}[n] * h[n] \\ &= \sum_{k=0}^{\mu} h[k]\tilde{x}[n-k] \\ &= \sum_{k=0}^{\mu} h[k]x[n-k]_N \\ &= x[n] \circledast h[n], \end{aligned} \tag{12.20}$$

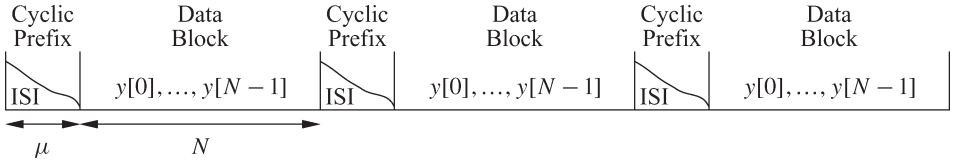


Figure 12.6: ISI between data blocks in channel output.

where the third equality follows from the fact that, for $0 \leq k \leq \mu$, $\tilde{x}[n-k] = x[n-k]_N$ for $0 \leq n \leq N-1$. Thus, by appending a cyclic prefix to the channel input, the linear convolution associated with the channel impulse response $y[n]$ for $0 \leq n \leq N-1$ becomes a circular convolution. Taking the DFT of the channel output in the absence of noise then yields

$$Y[i] = \text{DFT}\{y[n] = x[n] \circledast h[n]\} = X[i]H[i], \quad 0 \leq i \leq N-1, \tag{12.21}$$

and the input sequence $x[n], 0 \leq n \leq N-1$, can be recovered from the channel output $y[n], 0 \leq n \leq N-1$, for known $h[n]$ by

$$x[n] = \text{IDFT}\left\{\frac{Y[i]}{H[i]}\right\} = \text{IDFT}\left\{\frac{\text{DFT}\{y[n]\}}{\text{DFT}\{h[n]\}}\right\}. \tag{12.22}$$

Note that $y[n]$, $-\mu \leq n \leq N-1$, has length $N+\mu$, yet from (12.22) the first μ samples $y[-\mu], \ldots, y[-1]$ are not needed to recover $x[n], 0 \leq n \leq N-1$, owing to the redundancy associated with the cyclic prefix. Moreover, if we assume that the input $x[n]$ is divided into data blocks of size N with a cyclic prefix appended to each block to form $\tilde{x}[n]$, then the first μ samples of $y[n] = h[n] * \tilde{x}[n]$ in a given block are corrupted by ISI associated with the last μ samples of $x[n]$ in the prior block, as illustrated in Figure 12.6. The cyclic prefix serves to eliminate ISI between the data blocks, because the first μ samples of the channel output affected by this ISI can be discarded without any loss relative to the original information sequence. In continuous time this is equivalent to using a guard band of duration T_m (the channel delay spread) after every block of N symbols of duration NT_s in order to eliminate the ISI between these data blocks.

The benefits of adding a cyclic prefix come at a cost. Since μ symbols are added to the input data blocks, there is an overhead of μ/N and a resulting data-rate reduction of $N/(\mu+N)$. The transmit power associated with sending the cyclic prefix is also wasted because this prefix consists of redundant data. It is clear from Figure

12.6 that any prefix of length μ appended to input blocks of size N eliminates ISI between data blocks if the first μ samples of the block are discarded. In particular, the prefix can consist of all zero symbols, in which case – although the data rate is still reduced by $N/(N+\mu)$ – no power is used in transmitting the prefix. Trade-offs associated with the cyclic prefix versus this all-zero prefix, which is a form of vector coding, are discussed in Section 12.4.5.

The foregoing analysis motivates the design of OFDM. In OFDM the input data is divided into blocks of size N, where each block is referred to as an *OFDM symbol*. A cyclic prefix is added to each OFDM symbol to induce circular convolution of the input and channel impulse response. At the receiver, the output samples affected by ISI between OFDM symbols are removed. The DFT of the remaining samples are used to recover the original input sequence. The details of this OFDM system design are given in the next section.

Example 12.4: Consider an OFDM system with total bandwidth $B = 1$ MHz assuming $\beta = \varepsilon = 0$. A single-carrier system would have symbol time $T_s = 1/B = 1\,\mu$s. The channel has a maximum delay spread of $T_m = 5\,\mu$s, so with $T_s = 1\,\mu$s and $T_m = 5\,\mu$s there would clearly be severe ISI. Assume an OFDM system with MQAM modulation applied to each subchannel. To keep the overhead small, the OFDM system uses $N = 128$ subcarriers to mitigate ISI. So $T_N = NT_s = 128\,\mu$s. The length of the cyclic prefix is set to $\mu = 8 > T_m/T_s$ to ensure no ISI between OFDM symbols. For these parameters, find the subchannel bandwidth, the total transmission time associated with each OFDM symbol, the overhead of the cyclic prefix, and the data rate of the system assuming $M = 16$.

Solution: The subchannel bandwidth $B_N = 1/T_N = 7.812$ kHz, so $B_N \ll B_c = 1/T_m = 200$ kHz, ensuring negligible ISI. The total transmission time for each OFDM symbol is $T = T_N + \mu T_S = 128 + 8 = 136\,\mu$s. The overhead associated with the cyclic prefix is 8/128 which is 6.25%. The system transmits $\log_2 16 = 4$ bits per subcarrier every T seconds, so the data rate is $128 \cdot 4/136 \cdot 10^{-6} = 3.76$ Mbps, which is slightly less than $4B$ as a result of the cyclic prefix overhead.

12.4.3 Orthogonal Frequency-Division Multiplexing (OFDM)

The OFDM implementation of multicarrier modulation is shown in Figure 12.7. The input data stream is modulated by a QAM modulator, resulting in a complex symbol stream $X[0], X[1], \ldots, X[N-1]$. This symbol stream is passed through a serial-to-parallel converter, whose output is a set of N parallel QAM symbols $X[0], \ldots, X[N-1]$ corresponding to the symbols transmitted over each of the subcarriers. Thus, the N symbols output from the serial-to-parallel converter are the discrete frequency components of the OFDM modulator output $s(t)$. In order to generate $s(t)$, the frequency components are converted into time samples by performing an inverse DFT on these N symbols, which is efficiently implemented using the IFFT algorithm. The IFFT yields the OFDM symbol consisting of the sequence $x[n] = x[0], \ldots, x[N-1]$ of length N, where

$$x[n] = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X[i] e^{j2\pi ni/N}, \quad 0 \le n \le N-1. \tag{12.23}$$

This sequence corresponds to samples of the multicarrier signal: the multicarrier signal consists linearly modulated subchannels, and the right-hand side of (12.23) corresponds to samples of a sum of QAM symbols $X[i]$ each modulated by the carrier $e^{j2\pi it/T_N}$, $i = 0, \ldots, N-1$. The cyclic prefix is then added to the OFDM symbol, and the resulting time samples $\tilde{x}[n] = \tilde{x}[-\mu], \ldots, \tilde{x}[N-1] = x[N-\mu], \ldots, x[0], \ldots, x[N-1]$ are ordered by the parallel-to-serial converter and passed through a D/A converter, resulting after pulse shaping in the baseband OFDM signal $\tilde{x}(t)$, which is then upconverted to frequency f_0.

The transmitted signal is filtered by the channel impulse response and corrupted by additive noise, resulting in the received signal $r(t)$. This signal is downconverted to baseband and match-filtered to remove the high-frequency components. The A/D converter samples the resulting signal to obtain $y[n] = \tilde{x}[n] * h[n] + v[n]$, $-\mu \leq n \leq N-1$, where $h[n]$ is the discrete-time equivalent lowpass impulse response of the channel. The prefix of $y[n]$ consisting of the first μ samples is then removed. This results in N time samples whose DFT in the absence of noise is $Y[i] = H[i]X[i]$. These time samples are serial-to-parallel converted and passed through an FFT. This results in scaled versions of the original symbols $H[i]X[i]$, where $H[i] = H(f_i)$ is the flat fading channel gain associated with the ith subchannel. The FFT output is parallel-to-serial converted and passed through a QAM demodulator to recover the original data.

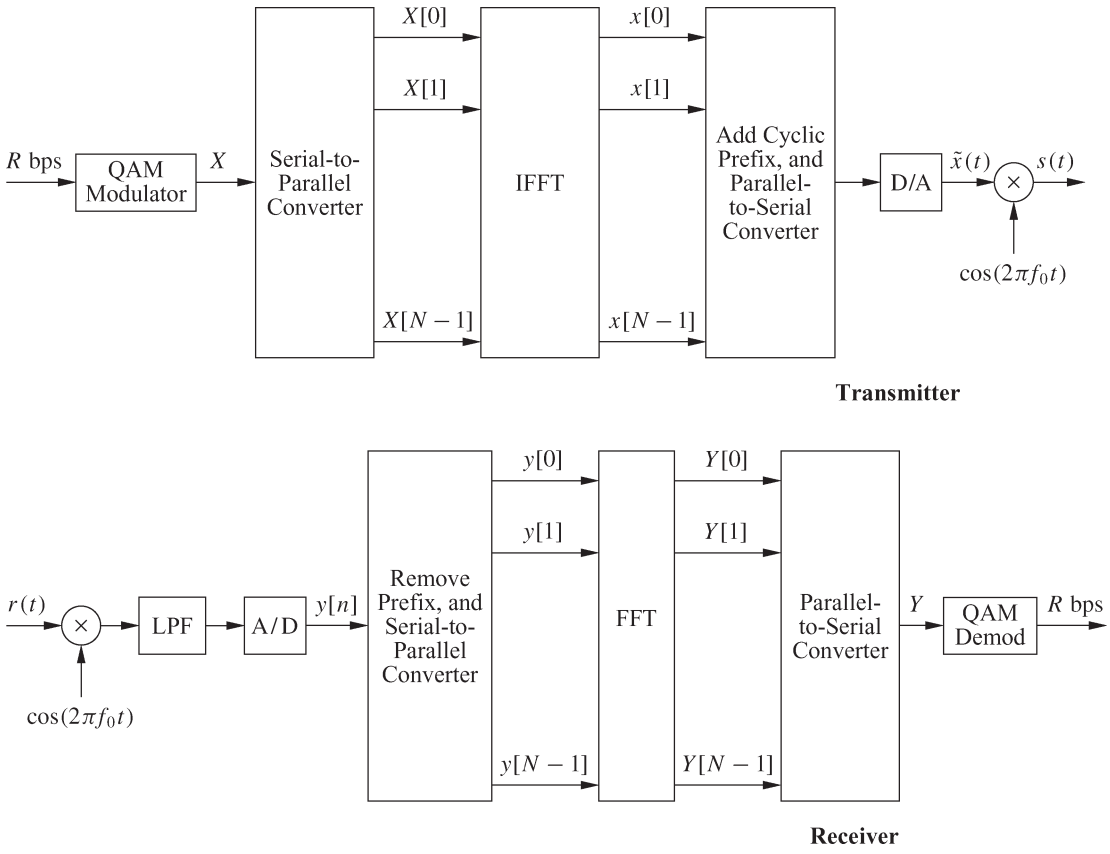


Figure 12.7: OFDM with IFFT/FFT implementation.

The OFDM system effectively decomposes the wideband channel into a set of narrowband orthogonal subchannels with a different QAM symbol sent over each subchannel. Knowledge of the channel gains $H[i]$, $i = 0, \ldots, N-1$, is not needed for this decomposition, in the same way that a continuous-time channel with frequency response $H(f)$ can be divided into orthogonal subchannels without knowledge of $H(f)$ by splitting the total signal bandwidth into nonoverlapping subbands. The demodulator can use the channel gains to recover the original QAM symbols by dividing out these gains: $X[i] = Y[i]/H[i]$. This process is called *frequency equalization*. However, as discussed in Section 12.3.2 for continuous-time OFDM, frequency equalization leads to noise enhancement because the noise in the ith subchannel is also scaled by $1/H[i]$. Hence, while the effect of flat fading on $X[i]$ is removed by this equalization, its received SNR is unchanged. Precoding, adaptive loading, and coding across subchannels (as discussed in Section 12.3) are better approaches to mitigate the effects of flat fading across subcarriers. An alternative to using the cyclic prefix is to use a prefix consisting of all zero symbols. In this case

the OFDM symbol consisting of $x[n]$, $0 \leq n \leq N-1$, is preceded by μ null samples, as illustrated in Figure 12.8. At the receiver the "tail" of the ISI associated with the end of a given OFDM symbol is added back in to the beginning of the symbol, which re-creates the effect of a cyclic prefix, so the rest of the OFDM system functions as usual. This zero prefix reduces the transmit power relative to a cyclic prefix by $N/(\mu+N)$, since the prefix does not require any transmit power. However, the noise from the received tail is added back into the beginning of the symbol, which increases the noise power by $(N+\mu)/N$. Thus, the difference in SNR is not significant for the two prefixes.

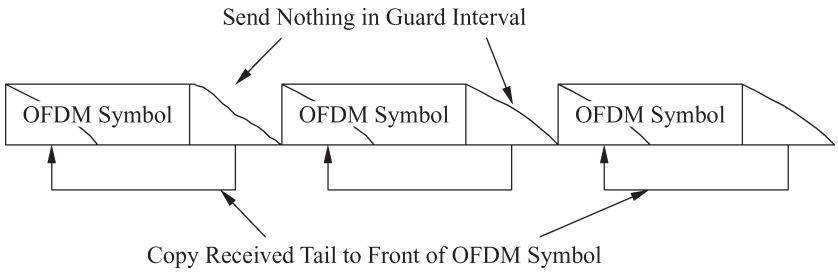


Figure 12.8: Creating a circular channel with an all-zero prefix.

12.4.4 Matrix Representation of OFDM

An alternate analysis for OFDM is based on a matrix representation of the system. Consider a discrete-time equivalent lowpass channel with FIR $h[n], 0 \leq n \leq \mu$, input $\tilde{x}[n]$, noise $v[n]$, and output $y[n] = \tilde{x}[n] * h[n] + v[n]$. Denote the nth element of these sequences as $h_n = h[n], \tilde{x}_n = \tilde{x}[n], v_n = v[n]$, and $y_n = y[n]$. With this notation the channel output sequence can be written in matrix form as

$$\begin{bmatrix} y_{N-1} \\ y_{N-2} \\ \vdots \\ y_0 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\mu & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{\mu-1} & h_\mu & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & \cdots & h_{\mu-1} & h_\mu \end{bmatrix} \begin{bmatrix} x_{N-1} \\ \vdots \\ x_0 \\ x_{-1} \\ \vdots \\ x_{-\mu} \end{bmatrix} + \begin{bmatrix} \nu_{N-1} \\ \nu_{N-2} \\ \vdots \\ \nu_0 \end{bmatrix}, \tag{12.24}$$

which can be written more compactly as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \nu. \tag{12.25}$$

The received symbols $y_{-1}, \ldots, y_{-\mu}$ are discarded because they are affected by ISI in the prior data block and are not needed to recover the input. The last μ symbols of $x[n]$ correspond to the cyclic prefix: $x_{-1} = x_{N-1}, x_{-2} = x_{N-2}, \ldots, x_{-\mu} = x_{N-\mu}$. From this it can be shown that the matrix representation (12.24) is equivalent to

$$\begin{bmatrix} y_{N-1} \\ y_{N-2} \\ \vdots \\ \vdots \\ \vdots \\ y_0 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\mu & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{\mu-1} & h_\mu & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & \cdots & h_{\mu-1} & h_\mu \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_2 & h_3 & \cdots & h_{\mu-2} & \cdots & h_0 & h_1 \\ h_1 & h_2 & \cdots & h_{\mu-1} & \cdots & 0 & h_0 \end{bmatrix} \begin{bmatrix} x_{N-1} \\ x_{N-2} \\ \vdots \\ \vdots \\ x_0 \end{bmatrix} + \begin{bmatrix} \nu_{N-1} \\ \nu_{N-2} \\ \vdots \\ \vdots \\ \vdots \\ \nu_0 \end{bmatrix}, \tag{12.26}$$

which can be written more compactly as

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \boldsymbol{\nu}. \tag{12.27}$$

This equivalent model shows that the inserted cyclic prefix allows the channel to be modeled as a circulant convolution matrix $\tilde{\mathbf{H}}$ over the N samples of interest. The matrix $\tilde{\mathbf{H}}$ is *normal* ($\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} = \tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$), so it has an eigenvalue decomposition

$$\tilde{\mathbf{H}} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^H, \tag{12.28}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues of $\tilde{\mathbf{H}}$ and $\mathbf{M}$ is a unitary matrix whose columns constitute the eigenvectors of $\tilde{\mathbf{H}}$.

It is straightforward to show that the DFT operation on $x[n]$ can be represented by the matrix multiplication

$$\mathbf{X} = \mathbf{Q}\mathbf{x}, \tag{12.29}$$

where $\mathbf{X} = (X[0], \ldots, X[N-1])^T$, $\mathbf{x} = (x[0], \ldots, x[N-1])^T$, and $\mathbf{Q}$ is an $N \times N$ matrix given by

$$\mathbf{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)^2} \end{bmatrix} \tag{12.30}$$

for $W_N = e^{-j2\pi/N}$. Since

$$\mathbf{Q}^{-1} = \mathbf{Q}^H, \tag{12.31}$$

the IDFT can be similarly represented as

$$\mathbf{x} = \mathbf{Q}^{-1}\mathbf{X} = \mathbf{Q}^H\mathbf{X}. \tag{12.32}$$

Let $\mathbf{v}$ be an eigenvector of $\tilde{\mathbf{H}}$ with eigenvalue λ. Then

$$\lambda\mathbf{v} = \tilde{\mathbf{H}}\mathbf{v}. \tag{12.33}$$

The unitary matrix $\mathbf{M}$ has columns that are the eigenvectors of $\tilde{\mathbf{H}}$; that is, $\lambda_i\mathbf{m}_i = \tilde{\mathbf{H}}\mathbf{m}_i$ for $i = 0, 1, \ldots, N-1$, where $\mathbf{m}_i$ denotes the ith column of $\mathbf{M}$. It can also be shown by induction that the columns of the DFT matrix $\mathbf{Q}^H$ are eigenvectors of $\tilde{\mathbf{H}}$, which implies that $\mathbf{Q} = \mathbf{M}^H$ and $\mathbf{Q}^H = \mathbf{M}$. Thus we have that

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{Q}\mathbf{y} \\
&= \mathbf{Q}[\tilde{\mathbf{H}}\mathbf{x} + \boldsymbol{\nu}] \\
&= \mathbf{Q}[\tilde{\mathbf{H}}\mathbf{Q}^H\mathbf{X} + \boldsymbol{\nu}] \\
&= \mathbf{Q}[\mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^H\mathbf{Q}^H\mathbf{X} + \boldsymbol{\nu}] \\
&= \mathbf{Q}\mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^H\mathbf{Q}^H\mathbf{X} + \mathbf{Q}\boldsymbol{\nu} \\
&= \mathbf{M}^H\mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^H\mathbf{M}\mathbf{X} + \mathbf{Q}\boldsymbol{\nu} \\
&= \boldsymbol{\Lambda}\mathbf{X} + \boldsymbol{\nu}_{\mathbf{Q}};
\end{aligned} \tag{12.34}
$$

here, since $\mathbf{Q}$ is unitary, it follows that $\boldsymbol{\nu}_{\mathbf{Q}} = \mathbf{Q}\boldsymbol{\nu}$ has the same noise autocorrelation matrix as $\boldsymbol{\nu}$ and hence is generally white and Gaussian, with unchanged noise power. Thus, this matrix analysis also shows that, by adding a cyclic prefix and using the IDFT/ DFT, OFDM decomposes an ISI channel into N orthogonal subchannels, and knowledge of the channel matrix $\mathbf{H}$ is not needed for this decomposition.

The matrix representation is also useful in analyzing OFDM systems with multiple antennas [29]. As discussed in Chapter 10, a MIMO channel is typically represented by an $M_r \times M_t$ matrix, where M_t is the number of transmit antennas and M_r the number of receive antennas. Thus, an MIMO-OFDM channel with N subchannels, M_t transmit antennas, M_r receive antennas, and a channel FIR of duration μ can be represented as

$$
\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\nu}, \tag{12.35}
$$

where $\mathbf{y}$ is a vector of dimension $M_r N \times 1$ corresponding to N output time samples at each of the M_r antennas, H is an $NM_r \times (N+\mu)M_t$ matrix corresponding to the N flat fading subchannel gains on each transmit-receive antenna pair, $\mathbf{x}$ is a vector of dimension $M_t(N+\mu) \times 1$ corresponding to N input time samples with appended cyclic prefix of length μ at each of the M_t transmit antennas, and ν is the channel noise vector of dimension $M_r N \times 1$. The matrix is in the same form as in the case of OFDM without multiple antennas, so the same design and analysis applies: with MIMO-OFDM the ISI is removed by breaking the wideband channel into many narrowband subchannels. Each subchannel experiences flat fading and so can be treated as a flat fading MIMO channel. The capacity of this channel is obtained by applying the same matrix analysis as for standard MIMO to the augmented channel with MIMO and OFDM [30]. In discrete implementations the input associated with each transmit antenna is broken into blocks of size N, with a cyclic prefix appended for converting linear convolution to circular and eliminating ISI between input blocks. More details can be found in [31].

12.4.5 Vector Coding

In OFDM the $N \times N$ circulant channel matrix $\tilde{\mathbf{H}}$ is decomposed using its eigenvalues and eigenvectors. Vector coding (VC) is a similar technique whereby the original $N \times (N+\mu)$ channel matrix $\mathbf{H}$ from (12.25) is decomposed using a singular value decomposition, which can be applied to a matrix of any dimension. The SVD does not require a cyclic prefix to make the subchannels orthogonal, so it is more efficient than OFDM in terms of energy. However, it is more complex and requires knowledge of the channel impulse response for the decomposition – in contrast to OFDM, which does not require channel knowledge for its decomposition.

From Appendix C, the singular value decomposition of $\mathbf{H}$ can be written as

$$
\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \tag{12.36}
$$

where $\mathbf{U}$ is $N \times N$ unitary, $\mathbf{V}$ is $(N+\mu) \times (N+\mu)$ unitary, and $\boldsymbol{\Sigma}$ is a matrix with ith diagonal element σ_i equal to the ith singular value of $\mathbf{H}$ and with zeros everywhere else. The singular values of $\mathbf{H}$ are related to the

eigenvalues of $\mathbf{H}\mathbf{H}^H$ by $\sigma_i = \sqrt{\lambda_i}$ for λ_i the ith eigenvalue of the matrix $\mathbf{H}\mathbf{H}^H$. Because $\mathbf{H}$ is a block-diagonal convolutional matrix, the rank of $\mathbf{H}$ is N and so $\sigma_i \neq 0$ for all $i = 1, \ldots, N$.

In vector coding, as in OFDM, input data symbols are grouped into vectors of N symbols. Let X_i, $i = 0, \ldots, N-1$, denote the symbol to be transmitted over the ith subchannel and let $\mathbf{X} = (X_0, \ldots, X_{N-1}, X_N, \ldots, X_{N+\mu-1})$ denote a vector of these symbols with μ symbols $X_N, \ldots, X_{N+\mu-1}$ (typically zeros) appended at the end. The data symbol X_i is multiplied by the ith column of $\mathbf{V}$ to form a vector, and then these vectors are added together. At the receiver, the received vector $\mathbf{y}$ is multiplied by each row of $\mathbf{U}^H$ to yield N output symbols Y_i, $i = 0, 1, \ldots, N-1$. This process is illustrated in Figure 12.9, where the multiplication by $\mathbf{V}$ and $\mathbf{U}^H$ performs a similar function as the transmit precoding and receiver shaping in MIMO systems.

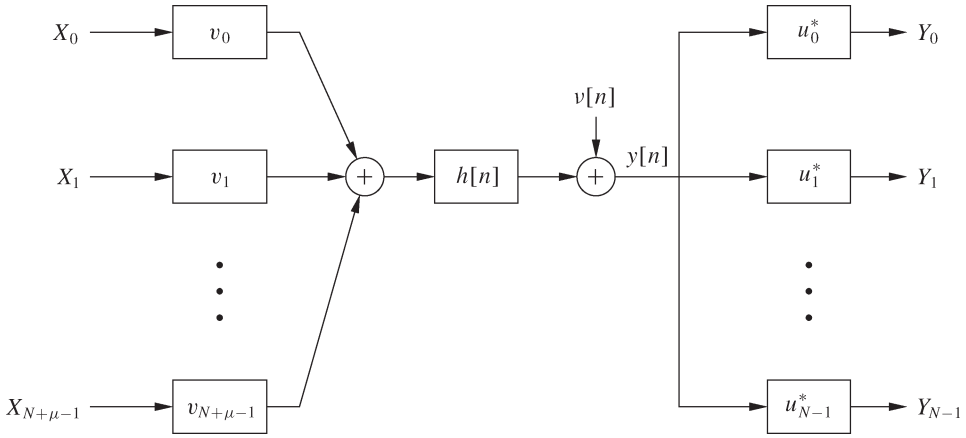


Figure 12.9: Vector coding.

Mathematically, it can be seen that the filtered transmit and received vectors are

$$\mathbf{x} = \mathbf{V}\mathbf{X} \tag{12.37}$$

and

$$\mathbf{Y} = \mathbf{U}^H\mathbf{y}. \tag{12.38}$$

As a result, it can be shown through simple linear algebra that the filtered received vector $\mathbf{Y}$ is ISI-free, since

$$\begin{aligned}\mathbf{Y} &= \mathbf{U}^H\mathbf{y} \\ &= \mathbf{U}^H(\mathbf{H}\mathbf{x} + \boldsymbol{\nu}) \\ &= \mathbf{U}^H(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H)\mathbf{V}\mathbf{X} + \mathbf{U}^H\boldsymbol{\nu} \\ &= \boldsymbol{\Sigma}\mathbf{X} + \mathbf{U}^H\boldsymbol{\nu}. \end{aligned} \tag{12.39}$$

Hence, each element of $\mathbf{X}$ is effectively passed through a scalar channel without ISI, where the scalar gain of subchannel i is the ith singular value of $\mathbf{H}$. Additionally, the new noise vector $\tilde{\boldsymbol{\nu}} = \mathbf{U}^H\boldsymbol{\nu}$ has unchanged noise variance, since $\mathbf{U}$ is unitary. The resulting received vector is thus

$$\begin{bmatrix} Y_{N-1} \\ Y_{N-2} \\ \vdots \\ Y_0 \end{bmatrix} = \begin{bmatrix} \sigma_1 X_{N-1} \\ \sigma_2 X_{N-2} \\ \vdots \\ \sigma_N X_0 \end{bmatrix} + \begin{bmatrix} \tilde{\nu}_{N-1} \\ \tilde{\nu}_{N-2} \\ \vdots \\ \tilde{\nu}_0 \end{bmatrix}. \tag{12.40}$$

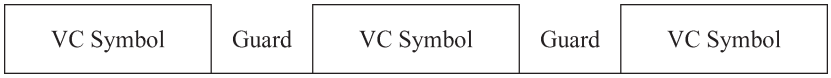


Figure 12.10: Guard interval (null prefix) in vector coding.

We call $\mathbf{X}$, the vector obtained by appending μ extra symbols to each block of N data symbols, a *vector codeword*. In contrast to OFDM, we see from (12.39) and the structure of $\mathbf{\Sigma}$ that the SVD does not require any particular value for these extra symbols; they need not be a cyclic prefix, nor must the "tail" be added back in if these symbols are all zeros. In practice the extra symbols are set to zero to save transmit power, thereby forming a guard band or null prefix between the vector codeword (VC) symbols, as shown in Figure 12.10.

Information and estimation theory have proven that vector coding is the optimal partition of the N-dimensional channel $\mathbf{H}$ [13, Chap. 4]. Thus, the capacity of any other channel partitioning scheme will be upper bounded by vector coding. Despite its theoretical optimality and ability to create ISI-free channels with relatively small overhead and no wasted transmit power, there are a number of important practical problems with vector coding. The two most important problems are as follows.

1. *Complexity*. With vector coding, as in simple multichannel modulation, the complexity scales quickly with N, the number of subcarriers. As seen from Figure 12.9, N transmit precoding and N receive shaping filters are required to implement vector coding. Furthermore, the complexity of finding the singular value decomposition of the $N \times (N + \mu)$ matrix $\mathbf{H}$ increases rapidly with N.

2. *SVD and channel knowledge*. In order to orthogonally partition the channel, the SVD of the channel matrix $\mathbf{H}$ must be computed. In particular, the precoding filter matrix must be known to the transmitter. This means that, every time the channel changes, a new SVD must be computed and the results conveyed to the transmitter. Generally, the computational complexity of the SVD and the delay incurred in getting the channel information back to the transmitter is prohibitive in wireless systems. Since OFDM can perform this decomposition without channel knowledge, OFDM is the method of choice for discrete multicarrier modulation in wireless applications.

Example 12.5: Consider a simple two-tap discrete-time channel (i.e. $\mu = 1$) described as

$$H(z) = 1 + .9z^{-1}. \tag{12.41}$$

Since $\mu + 1 = T_m/T_s = 2$, with $N = 8$ we ensure $B_N \approx 1/(NT_s) \ll B_c \approx 1/T_m$. Find the system matrix representation (12.25) and the singular values of the associated channel matrix $\mathbf{H}$.
Solution: The representation (12.25) for $H(z) = 1 + .9z^{-1}$ and $N = 8$ is given by

$$\begin{bmatrix} y_7 \\ y_6 \\ \vdots \\ y_0 \end{bmatrix} = \begin{bmatrix} 1 & .9 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & .9 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 1 & .9 \end{bmatrix} \begin{bmatrix} x_7 \\ x_6 \\ \vdots \\ x_{-1} \end{bmatrix} + \begin{bmatrix} \nu_7 \\ \nu_6 \\ \vdots \\ \nu_0 \end{bmatrix}. \tag{12.42}$$

The singular values of the matrix $\mathbf{H}$ in (12.42) can be found via a standard computer package (e.g. Matlab) as

$$\mathbf{\Sigma} = \text{diag}[1.87, 1.78, 1.65, 1.46, 1.22, .95, .66, .34]. \tag{12.43}$$

The precoding and combining matrices $\mathbf{U}$ and $\mathbf{V}$ are also easily found. Given $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Sigma}$, this communication is ISI-free, with the symbols $X_0, X_1, \ldots, X_{N-1}$ being multiplied by the corresponding singular values as in (12.40).

12.5 Challenges in Multicarrier Systems

12.5.1 Peak-to-Average Power Ratio

The peak-to-average power ratio (PAR) is an important attribute of a communication system. A low PAR allows the transmit power amplifier to operate efficiently, whereas a high PAR forces the transmit power amplifier to have a large *backoff* in order to ensure linear amplification of the signal. This is demonstrated in Figure 12.11, which shows a typical power amplifier response. Operation in the linear region of this response is generally required to avoid signal distortion, so the peak value is constrained to be in this region. Clearly it would be desirable to have the average and peak values be as close together as possible in order for the power amplifier to operate at maximum efficiency. Additionally, a high PAR requires high resolution for the receiver A/D converter, since the dynamic range of the signal is much larger for high-PAR signals. High-resolution A/D conversion places a complexity and power burden on the receiver front end.

The PAR of a continuous-time signal is given by

$$\text{PAR} \triangleq \frac{\max_t |x(t)|^2}{\mathbf{E}_t[|x(t)|^2]} \tag{12.44}$$

and for a discrete-time signal is given by

$$\text{PAR} \triangleq \frac{\max_n |x[n]|^2}{\mathbf{E}_n[|x[n]|^2]}. \tag{12.45}$$

Any constant amplitude signal (e.g., a square wave) has $\text{PAR} = 0$ dB. A sine wave has $\text{PAR} = 3$ dB because $\max[\sin^2(2\pi t/T)] = 1$ and

$$\mathbf{E}\left[\sin^2\left(\frac{2\pi t}{T}\right)\right] = \frac{1}{T}\int_0^T \sin^2\left(\frac{2\pi t}{T}\right) dt = .5, \tag{12.46}$$

so PAR $= 1/.5 = 2$.

In general, PAR should be measured with respect to the continuous-time signal using (12.44), since the input to the amplifier is an analog signal. The PAR given by (12.44) is sensitive to the pulse shape $g(t)$ used in the modulation, and it does not generally lead to simple analytical formulas [32]. For illustration we will focus on the PAR associated with the discrete-time signal, since it lends itself to a simple characterization. However, care must be taken when interpreting these results, since they can be quite inaccurate if the pulse shape $g(t)$ is not taken into account.

Consider the time-domain samples that are output from the IFFT:

$$x[n] = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X[i] e^{j2\pi in/N}, \quad 0 \le n \le N-1. \tag{12.47}$$

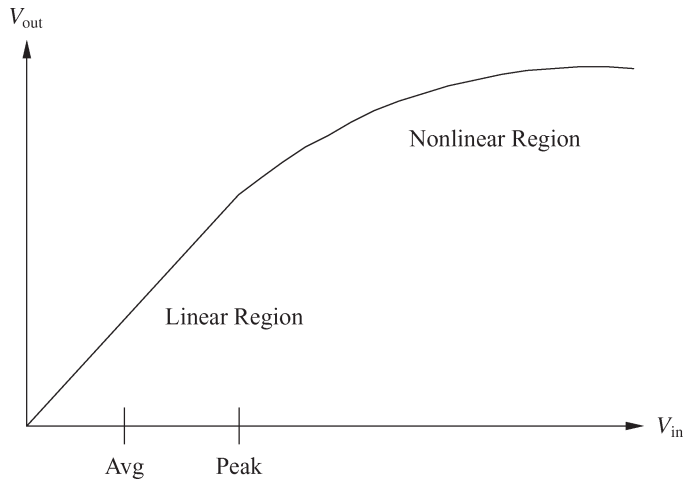


Figure 12.11: Typical power amplifier response.

If N is large then the central limit theorem is applicable, and $x[n]$ are zero-mean complex Gaussian random variables because the real and imaginary parts are summed. The Gaussian approximation for IFFT outputs is generally quite accurate for a reasonably large number of subcarriers ($N \geq 64$). For $x[n]$ complex Gaussian, the envelope of the OFDM signal is Rayleigh distributed with variance σ^2, and the phase of the signal is uniform. Since the Rayleigh distribution has infinite support, the peak value of the signal will exceed any given value with nonzero probability. It can then be shown [33] that the probability that the PAR given by (12.45) exceeds a threshold $P_0 = \sigma_0^2/\sigma^2$ is given by

$$p(\text{PAR} \geq P_0) = 1 - (1 - e^{-P_0})^N. \tag{12.48}$$

Let us now investigate how PAR grows with the number of subcarriers. Consider N Gaussian independent and identically distributed random variables $x_n (0 \leq n \leq N-1)$ with zero mean and unit power. The average signal power $\mathbf{E}_n[|x[n]|^2]$ is then

$$\begin{aligned} \mathbf{E}\left[\frac{1}{\sqrt{N}}|x_0 + x_1 + \cdots + x_{N-1}|^2\right] &= \frac{1}{N}\mathbf{E}|x_0 + x_1 + \cdots + x_{N-1}|^2 \\ &= \frac{\mathbf{E}|x_0^2| + \mathbf{E}|x_1^2| + \cdots + \mathbf{E}|x_{N-1}^2|}{N} \\ &= 1. \end{aligned} \tag{12.49}$$

The maximum value occurs when all the x_i add coherently, in which case

$$\max\left[\frac{1}{\sqrt{N}}|x_0 + x_1 + \cdots + x_{N-1}|\right]^2 = \left|\frac{N}{\sqrt{N}}\right|^2 = N. \tag{12.50}$$

Hence the maximum PAR is N for N subcarriers. In practice, full coherent addition of all N symbols is highly improbable and so the observed PAR is typically less than N – usually by many decibels. Nevertheless, PAR increases approximately linearly with the number of subcarriers. So even though it is desirable to have N as large as possible in order to keep the overhead associated with the cyclic prefix down, a large PAR is an important penalty that must be paid for large N.

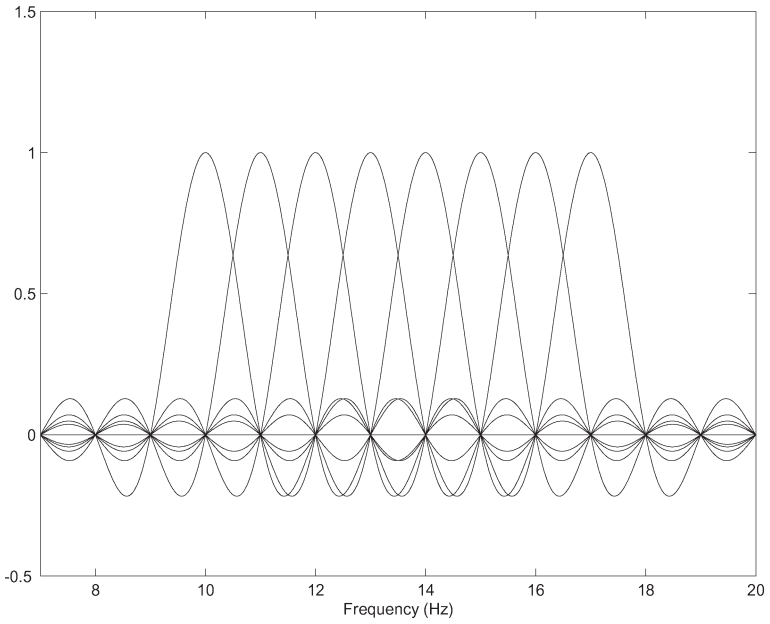


Figure 12.12: OFDM overlapping subcarriers for rectangular pulses with $f_0 = 10$ Hz and $\triangle f = 1$ Hz.

There are a number of ways to reduce or tolerate the PAR of OFDM signals, including clipping the OFDM signal above some threshold, peak cancellation with a complementary signal, allowing nonlinear distortion from the power amplifier (and correction for it), and special coding techniques [34]. A good summary of some of these methods can be found in [35].

12.5.2 Frequency and Timing Offset

We have seen that OFDM modulation encodes the data symbols X_i onto orthogonal subchannels, where orthogonality is assured by the subcarrier separation $\triangle f = 1/T_N$. The subchannels may overlap in the frequency domain, as shown in Figure 12.12 for a rectangular pulse shape in time (sinc function in frequency). In practice, the frequency separation of the subcarriers is imperfect and so $\triangle f$ is not exactly equal to $1/T_N$. This is generally caused by mismatched oscillators, Doppler frequency shifts, or timing synchronization errors. For example, if the carrier frequency oscillator is accurate to 1 part per million then the frequency offset $\triangle f_\varepsilon \approx f_0(.1 \cdot 10^{-6})$. If $f_0 = 5$ GHz, the carrier frequency for 802.11a WLANs, then $\triangle f_\varepsilon = 500$ Hz, which will degrade the orthogonality of the subchannels because now the received samples of the FFT will contain interference from adjacent subchannels. Next we'll analyze this intercarrier interference (ICI) more closely.

The signal corresponding to subcarrier i can be simply expressed for the case of rectangular pulse shapes (suppressing the data symbol and the carrier frequency) as

$$x_i(t) = e^{j2\pi it/T_N}. \tag{12.51}$$

An interfering subchannel signal can be written as

$$x_{i+m}(t) = e^{j2\pi(i+m)t/T_N}. \tag{12.52}$$

If the signal is demodulated with a frequency offset of δ/T_N, then this interference becomes

$$x_{i+m}(t) = e^{j2\pi(i+m+\delta)t/T_N}. \tag{12.53}$$

The ICI between subchannel signals x_i and x_{i+m} is simply the inner product between them:

$$I_m = \int_0^{T_N} x_i(t) x_{i+m}^*(t)\, dt = \frac{T_N(1 - e^{-j2\pi(\delta+m)})}{j2\pi(m+\delta)}. \tag{12.54}$$

It can be seen that, in (12.54), $\delta = 0 \Rightarrow I_m = 0$ as expected. The total ICI power on subcarrier i is then

$$\mathrm{ICI}_i = \sum_{m \neq i} |I_m|^2 \approx C_0 (T_N \delta)^2, \tag{12.55}$$

where C_0 is some constant. Several important trends can be observed from this simple approximation. First, as T_N increases, the subcarriers grow narrower and hence more closely spaced, which then results in more ICI. Second, the ICI predictably grows with the frequency offset δ, and the growth is about quadratic. Another interesting observation is that (12.55) does not appear to be directly affected by N. But picking N large generally forces T_N to be large also, which then causes the subcarriers to be closer together. Along with the larger PAR that comes with large N, the increased ICI is another reason to pick N as low as possible, assuming the overhead budget can be met. In order to further reduce the ICI for a given choice of N, nonrectangular windows can also be used [36, 37].

The effects from timing offset are generally less than those from the frequency offset [5, Chap. 5], as long as a full N-sample OFDM symbol is used at the receiver without interference from the previous or subsequent OFDM symbols (this is ensured by taking the cyclic prefix length $\mu \gg \sigma_{T_m}/T_s$, where σ_{T_m} is the channel's rms delay spread). It can be shown that the ICI power on subcarrier i due to a receiver timing offset τ can be approximated as $2(\tau/T_N)^2$. Since usually $\tau \ll T_N$, this effect is typically negligible.

12.5.3 Large Doppler Spread

OFDM modulation is designed to compensate for delay spread. In channels that are constant over an OFDM block, the rate within each subchannel can be adapted or frequency-domain equalization used to equalize the subchannel gains. However, these techniques only work when the channel is roughly constant over an OFDM block, i.e. when the channel Doppler spread is small.

Recall that OFDM utilizes time-domain multiplexing across the different subcarriers as well as a cycle prefix in time to cancel ISI between OFDM blocks. For channels that have little delay spread and a large Doppler spread, a time-frequency dual of the OFDM strategy uses frequency-domain multiplexing across different timeslots with a frequency-domain cyclic prefix to remove Doppler spread. This technique can compensate for the effect of an arbitrary Doppler spread by performing linear equalization in the time domain, which is the dual of frequency-domain equalization performed in OFDM. A performance analysis of this technique can be found in [39].

Channels that have both a large delay spread and a large Doppler spread are referred to as *doubly-dispersive* channels since they exhibit dispersion in both the time and frequency domains. Such channels vary within a time-frequency OFDM block, thereby precluding accurate channel estimates. OFDM typically performs poorly in doubly-dispersive channels as, without accurate channel estimates, neither adaptive subcarrier modulation nor frequency-domain equalization can be used to mitigate subcarrier fading. One approach to mitigate this performance degradation is to limit the symbol time or the OFDM block duration. Indeed, it is typical for OFDM deployments in time-varying channels to limit the symbol time, or, equivalently, increase the subcarrier spacing, so that the product of the overall symbol duration and the Doppler spread is small. In this regime, the time-frequency dispersive channel behaves approximately as a static delay spread channel at the expense of an increased cyclic prefix overhead. A detailed performance analysis of such schemes is presented in [38]. Alternatively, multicarrier modulation and demodulation techniques can be designed based on the delay-Doppler characteristics of the channel rather than its time-frequency characteristics. The delay-Doppler characteristics of a channel vary far more slowly than its time-frequency characteristics. Thus, these technique can be designed to operate on a quasi-static channel. Multicarrier techniques for delay-Doppler channels are discussed in Section 12.7.

12.6 Case Study: OFDM design in the Wi-Fi Standard

The IEEE 802.11 Wi-Fi family of standard, which occupies 80 MHz of bandwidth in the 2.4 GHz unlicensed band and 300 MHz of bandwidth in the 5-GHz unlicensed band, is based on OFDM [6], except for the first standard, 802.11b, which uses spread spectrum. The IEEE 802.11b and 802.11g standards operated in the crowded 2.4-GHz band [7], 802.11a operates in the 5 GHz band, and all generations of Wi-Fi starting with 802.11n operate in both the 2.4GHz and 5GHz bands. In this section we study the properties of OFDM design in the Wi-Fi standard and discuss some of the design choices. For simplicity we will focus on the 802.11a standard in the 300 MHz band since it does not incorporate MIMO, in order to focus on the OFDM design choices in WiFi design.

In 802.11a, $N = 64$ subcarriers are generated, although only 48 are actually used for data transmission: the outer 12 are zeroed in order to reduce adjacent channel interference, and 4 are used as pilot symbols for channel estimation. The cyclic prefix consists of $\mu = 16$ samples, so the total number of samples associated with each OFDM symbol, including both data samples and the cyclic prefix, is 80. Coded bits are packetized, and the transmitter gets periodic feedback from the receiver about the packet error rate, which it uses to pick an appropriate error correction code and modulation technique. The same code and modulation must be used for *all* the subcarriers at any given time. The error correction code is a convolutional code with one of three possible code rates: $r = 1/2, 2/3$, or 3/4. The modulation types that can be used on the subchannels are BPSK, QPSK, 16-QAM, or 64-QAM.

The total system bandwidth of 300 MHz is divided into 20-MHz channels that can be assigned to different users. Since the channel bandwidth B (and sampling rate $1/T_s$) is 20 MHz and since there are 64 subcarriers evenly spaced over that bandwidth, the subcarrier bandwidth is:

$$B_N = \frac{20 \text{ MHz}}{64} = 312.5 \text{ kHz}. \tag{12.56}$$

Since $\mu = 16$ and $1/T_s = 20$ MHz, the maximum delay spread for which ISI is removed is

$$T_m < \mu T_s = \frac{16}{20 \text{ MHz}} = 0.8 \ \mu\text{s}, \tag{12.57}$$

which corresponds to delay spread in an indoor environment. Including both the OFDM symbol and cyclic prefix, there are $80 = 64 + 16$ samples per OFDM symbol time, so the symbol time per subchannel is

$$T_N = 80T_s = \frac{80}{20 \cdot 10^6} = 4 \ \mu\text{s}. \tag{12.58}$$

The data rate per subchannel is $\log_2 M/T_N$. Thus, the minimum data rate for this system – corresponding to BPSK (1 bit/symbol), an $r = 1/2$ code, and taking into account that only 48 subcarriers actually carry usable data – is given by

$$\begin{aligned} R_{\min} &= 48 \text{ subcarriers} \cdot \frac{1/2 \text{ bit}}{1 \text{ coded bit}} \cdot \frac{1 \text{ coded bit}}{\text{subcarrier symbol}} \cdot \frac{1 \text{ subcarrier symbol}}{4 \cdot 10^{-6} \text{ seconds}} \\ &= 6 \text{ Mbps}. \end{aligned} \tag{12.59}$$

The maximum data rate that can be transmitted is

$$\begin{aligned} R_{\max} &= 48 \text{ subcarriers} \cdot \frac{3/4 \text{ bit}}{1 \text{ coded bit}} \cdot \frac{6 \text{ coded bits}}{\text{subcarrier symbol}} \cdot \frac{1 \text{ subcarrier symbol}}{4 \cdot 10^{-6} \text{ seconds}} \\ &= 54 \text{ Mbps}. \end{aligned} \tag{12.60}$$

Naturally, a wide range of data rates between these two extremes is possible.

Example 12.6: Find the data rate of an 802.11a system assuming 16-QAM modulation and rate-2/3 coding.
Solution: With 16-QAM modulation, each subcarrier transmits $\log_2(16) = 4$ coded bits per subcarrier symbol and there are a total of 48 subcarriers used for data transmission. With a rate-2/3 code, each coded bit relays 2/3 of an information bit per T_N seconds. Thus, the data rate is given by

$$\begin{aligned} R_{\max} &= 48 \text{ subcarriers} \cdot \frac{2/3 \text{ bit}}{1 \text{ coded bit}} \cdot \frac{4 \text{ coded bits}}{\text{subcarrier symbol}} \cdot \frac{1 \text{ subcarrier symbol}}{4 \cdot 10^{-6} \text{ seconds}} \\ &= 32 \text{ Mbps.} \end{aligned} \tag{12.61}$$

12.7 Orthogonal Time-Frequency-Space Modulation (OTFS)

OFDM is designed in the time-frequency domain and exploits the properties of the FFT and IFFT for its simple implementation. The use of adaptive loading or frequency equalization leads to excellent performance of OFDM in channels whose coherence time is larger than an OFDM symbol time. However, as discussed in Section 12.5.3, for high Doppler channels where the channel coherence time is on the order of an OFDM symbol, adaptive loading and frequency domain equalization perform poorly, since accurate channel estimates cannot be obtained. Moreover, while diversity techniques can mitigate the effect of time or frequency selective fading without channel estimation, an OFDM symbol must span both the channel coherence time and the channel coherence bandwidth to experience the full diversity of the channel, which may be impractical for short data transmissions. These shortcomings of OFDM in high-Doppler channels have led to alternative multicarrier techniques based on signal processing in the delay-Doppler domain.

As discussed in Section 3.3.1, the scattering function $C(\tau, \rho)$, corresponding to the time-varying channel impulse response $c(\tau, t)$, captures the delay-Doppler characteristics of a wireless channel via the Doppler frequency parameter ρ and the delay parameter τ. An advantage of designing multicarrier techniques in the delay-Doppler versus the time-frequency domain is that the delay and Doppler characteristics for wireless channels correspond to those of the multipath components introduced by the physical propagation environment. In typical environments the number of multipath components and their delay and Doppler characteristics change much more slowly than the channel coherence time associated with the maximum Doppler frequency. Hence, a channel with high Doppler will be quasi-static in the delay-Doppler domain. Moreover, channels characterized by a small number N of multipath components can be described in the delay-Doppler domain with $3N$ parameters corresponding to the received amplitude, delay, and Doppler of each multipath component. Characterizing the same channel in the time-frequency domain requires the power level associated with each time-frequency block. Multicarrier methods designed in the delay-Doppler domain have been developed to exploit these advantages over OFDM [40, 41, 42], the most common of which is Orthogonal Time-Frequency-Space (OTFS) modulation [43, 44, 45], first proposed in the landmark paper [46].

Orthogonal Time-Frequency-Space signaling (OTFS) assigns modulated symbols to discrete blocks in the two-dimensional delay-Doppler domain. It then converts these modulated symbols first to the time-frequency domain and then to the time domain, as illustrated in Fig. 12.13. This transformation spreads each modulated symbol over a large time-frequency block. Thus each OTFS symbol benefits from both time and frequency diversity introduced by the channel. The reverse operations are performed at the receiver, also illustrated in Fig. 12.13. For

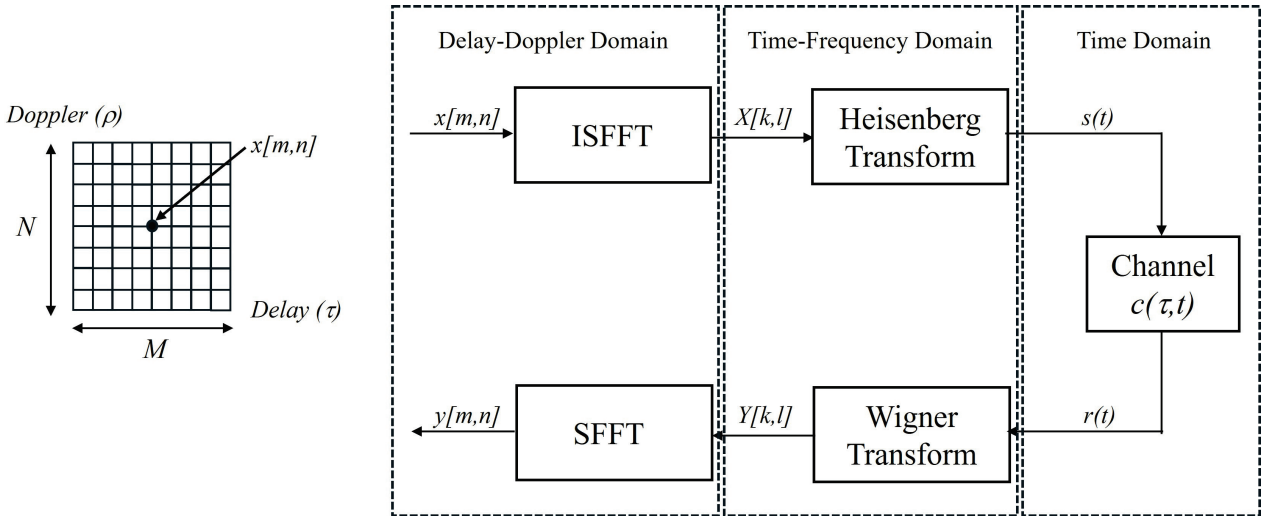


Figure 12.13: OTFS Modulation and Demodulation

ideal channels, each modulated symbol experiences the exact same channel gain, which is quasi-static due to the full channel diversity it captures.

The two-dimensional OTFS modulation and demodulation work as follows. The modulation first divides up the delay-Doppler domain of the channel scattering function $C(\tau, \rho)$ into MN blocks $[m\Delta\tau, n\Delta\rho], M = 0, \ldots, M-1; n = 0, \ldots, N-1$. Typically M and $\Delta\tau$ are chosen such that $M\Delta\tau$ is less than the channel delay spread. Similarly, N and $\Delta\rho$ are chosen such that $N\Delta\rho$ is less than the channel Doppler spread. Each two-dimensional OTFS symbol consists of MN QAM symbols $x[m,n], m = 1, \ldots, M-1; n = 1, \ldots, N-1$ where each QAM symbol is assigned to the associated delay-Doppler block $[m\Delta\tau, n\Delta\rho]$.

The OTFS symbol is converted to the discrete time-frequency domain using the two-dimensional inverse FFT, also called the inverse Symplectic Finite Fourier Transform (ISFFT), followed by windowing, as follows:

$$X[k,l] = W_{tx}[k,l]\text{ISFFT}(x[m,n]) = \frac{W_{tx}[k,l]}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m,n] e^{-j2\pi\left(\frac{lm}{M} - \frac{kn}{N}\right)}. \tag{12.62}$$

As discussed below, the transmit window function $W_{tx}[k,l]$ is used in the overall design to minimize cross-symbol interference. Note that for simplicity, this windowing operation is not shown in Fig. 12.13. The two-dimensional discrete-time representation of the OTFS symbol in time and frequency, $X[k,l]$, is then converted to the transmit signal $s(t)$ using the *Heisenberg transform*, given as follows for transmit pulse shape $g_s(t)$, frequency sampling interval Δf, and time sampling interval T:

$$s(t) = \sum_{l=-M/2}^{M/2-1} \sum_{k=0}^{N-1} X[k,l] g_s(t - kT) e^{j2\pi l\Delta f(t-kT)} \tag{12.63}$$

.

The reverse operations are performed at the receiver for demodulation: the received signal $r(t)$ is first converted to the time-frequency domain using the Wigner transform. This signal is multiplied by a receive window function (also not shown in Fig. 12.13 for simplicity), and then converted to the delay-Doppler domain using the Symplectic Finite Fourier Transform (SFFT). This yields the received symbols $y[m,n]$ which are demodulated symbol-by-symbol. Specifically, the Wigner transform performs matched filtering and time-frequency sampling on the received signal $r(t)$ to obtain the time-frequency symbols $Y[k,l]$ as samples of the *cross ambiguity function*

$A(\tau, \rho)$. This function, which is parameterized by receiver matched filter $g_r(t)$, frequency sampling interval Δf, and time sampling interval T, is defined as

$$A(\tau, \rho) = \int g_r(t - \tau) r(t) e^{-j2\pi\rho(t-\tau)} dt. \tag{12.64}$$

The time-frequency symbols $Y[k, l] = A(kT, l\Delta f)$ are obtained by sampling $A(\tau, \rho)$ at time intervals T and frequency intervals Δf. These time-frequency symbols are multiplied by the receive window function $W_{rx}[k, l]$ and then mapped back to the delay-Doppler domain using the SFFT as follows:

$$y[m, n] = \text{SFFT}(W_{rx}[k, l]Y[k, l]) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} W_{rx}[k, l]Y[k, l] e^{-j2\pi\left(\frac{kn}{N} - \frac{lm}{M}\right)}. \tag{12.65}$$

It is shown in [43] that $y[m, n]$ in the absence of noise equals the two-dimensional convolution of the transmitted symbol $x[m, n]$ with a sampled version of the windowed channel scattering function $C_w(k, l)$ as follows:

$$y[k, l] = \frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[m, n] C_w \left(\frac{k - m}{MT}, \frac{l - n}{N\Delta\rho} \right) \tag{12.66}$$

where the windowed channel scattering function $C_w(\rho, \tau)$ is given by

$$C_w(\tau, \rho) = \iint c(x, y) w(\tau - x, \rho - y) e^{-j2\pi xy} dy dx. \tag{12.67}$$

for time-varying channel impulse response $c(\tau, \rho)$ and receive windowing function $w_(x, y)$. This windowing function is the simplectic discrete Fourier transform of product of the transmit and receive windowing filters:

$$w(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W[m, n] e^{-j2\pi(xmT - \rho n\Delta f)} \tag{12.68}$$

for $W[n, m] = W_{tx}[n, m]W_{rx}[n, m]$. Ideally, these filters can be designed such that the samples of $c(k, l) = 0$ except for $c(0, 0)$. In this case, all transmitted symbols experience the full time and frequency diversity of the channel via the common channel gain $c(0, 0)$, and there is no interference between symbols. The channel fading gain $c(0, 0)$ can be mitigated using equalization or coding. In non-ideal scenarios, design of the transmit and receive pulse shapes and windowing functions based on the physical channel $C(\tau, \rho)$ aims to minimize interference between received symbols. Interference cancellation techniques can be employed at the receiver to remove any remaining interference [47]. The quasi-static nature of the delay-Doppler channel and the fact that all symbols experience the same channel gain often leads to significantly better performance of OTFS relative to OFDM in channels with a large Doppler spread [46, 43, 47]. A variant of OTFS - implemented using the Zak transform - allows for greater predictability of the delay-Doppler channel relative to the Heisenberg-Wigner transform approach, leading to better performance at high Doppler frequencies [48]. An in-depth discussion of this OTFS variant can be found in [49].

Chapter 12 Problems

1. Show that the minimum separation $\triangle f$ for subcarriers $\{\cos(2\pi j \triangle f t + \phi_j),\ j = 1, 2, \ldots\}$ to form a set of orthonormal basis functions on the interval $[0, T_N]$ is $\triangle f = 1/T_N$ for any initial phase ϕ_j. Show that if $\phi_j = 0$ for all j then this carrier separation can be reduced by half.

2. Consider an OFDM system operating in a channel with coherence bandwidth $B_c = 10$ kHz.

   (a) Find a subchannel symbol time $T_N = 1/B_N = 10T_m$, assuming $T_m = 1/B_c$. This should ensure flat fading on the subchannels.

   (b) Assume the system has $N = 128$ subchannels. If raised cosine pulses with $\beta = 1.5$ are used and if the required additional bandwidth (from time limiting) to ensure minimal power outside the signal bandwidth is $\varepsilon = .1$, then what is the total bandwidth of the system?

   (c) Find the total required bandwidth of the system using overlapping carriers separated by $1/T_N$, and compare with your answer in part (b).

3. Show from the definition of the DFT that circular convolution of discrete time sequences leads to multiplication of their DFTs.

4. Consider a data signal with a bandwidth of .5 MHz and a data rate of .5 Mbps. The signal is transmitted over a wireless channel with a delay spread of 10 μs.

   (a) If multicarrier modulation with nonoverlapping subchannels is used to mitigate the effects of ISI, approximately how many subcarriers are needed? What is the data rate and symbol time on each subcarrier? (We do not need to eliminate the ISI completely, so $T_s = T_m$ is sufficient for ISI mitigation.)

   Assume for the remainder of the problem that the average received SNR (γ_s) on the nth subcarrier is $1000/n$ (linear units) and that each subcarrier experiences flat Rayleigh fading (so ISI is completely eliminated).

   (b) Suppose BPSK modulation is used for each subcarrier. If a repetition code is used across all subcarriers (i.e., if a copy of each bit is sent over each subcarrier), then what is the BER after majority decoding? What is the data rate of the system?

   (c) Suppose you use adaptive loading (i.e., different constellations on each subcarrier) such that the average BER on each subcarrier does not exceed 10^{-3} (this is averaged over the fading distribution; do not assume that the transmitter and receiver adapt power or rate to the instantaneous fade values). Find the MQAM constellation that can be transmitted over each subcarrier while meeting this average BER target. What is the total data rate of the system with adaptive loading?

5. Consider a multicarrier modulation transmission scheme with three nonoverlapping subchannels spaced 200 kHz apart (from carrier to carrier) and with subchannel baseband bandwidth of 100 kHz.

   (a) For what values of the channel coherence bandwidth will the subchannels of your multicarrier scheme exhibit flat fading (approximately no ISI)? For which such values will the subcarriers exhibit independent fading? If the subcarriers exhibit correlated fading, what impact will this have on coding across subchannels?

   (b) Suppose that you have a total transmit power $P = 300$ mW and that the noise power in each subchannel is 1 mW. With equal power of 100 mW transmitted on each subchannel, the received SNR on each subchannel is $\gamma_1 = 11$ dB, $\gamma_2 = 14$ dB, and $\gamma_3 = 18$ dB. Assume the subchannels do not experience

fading, so these SNRs are constant. For these received SNRs, find the maximum signal constellation size for MQAM that can be transmitted over each subchannel for a target BER of 10^{-3}. Assume the MQAM constellation is restricted to be a power of 2 and use the bound BER $\leq .2e^{-1.5\gamma/(M-1)}$ for your calculations. Find the corresponding total data rate of the multicarrier signal assuming a symbol rate on each subchannel of $T_s = 1/B$, where B is the baseband subchannel bandwidth.

(c) For the subchannel SNRs given in part (b), suppose we want to use precoding (to equalize the received SNR in each subchannel) and then send the same signal constellation over each subchannel. What size signal constellation is needed to achieve the same data rate as in part (b)? What transmit power would be needed on each subchannel to achieve the required received SNR for this constellation with a 10^{-3} BER target? How much must the total transmit power be increased over the 300-mW transmit power in part (b)?

6. Consider a channel with impulse response

$$h(t) = \alpha_0\delta(t) + \alpha_1\delta(t - T_1) + \alpha_2\delta(t - T_2).$$

Assume that $T_1 = 10\,\mu$s and $T_2 = 20\,\mu$s. You want to design a multicarrier system for the channel with subchannel bandwidth $B_N = B_c/2$. If raised cosine pulses with $\beta = 1$ are used and if the subcarriers are separated by the minimum bandwidth necessary to remain orthogonal, then what is the total bandwidth occupied by a multicarrier system with eight subcarriers? Assuming a constant SNR on each subchannel of 20 dB, find the maximum constellation size for MQAM modulation that can be sent over each subchannel with a target BER of 10^{-3}, assuming also that M is restricted to be a power of 2. Find the corresponding total data rate of the system.

7. Show the equivalence of matrix representations (12.24) and (12.26) for OFDM when a cyclic prefix is appended to each OFDM symbol.

8. Show that the DFT operation on $x[n]$ can be represented by the matrix multiplication $X[i] = \mathbf{Q}x[n]$, where

$$\mathbf{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)^2} \end{bmatrix}$$

for $W_N = e^{-j2\pi/N}$.

9. This problem shows that the rows of the DFT matrix $\mathbf{Q}$ are eigenvectors of the circulant convolution matrix $\tilde{\mathbf{H}}$ defined in Eqn. 12.28.

(a) Show that the first row of $\mathbf{Q}$ is an eigenvector of $\tilde{\mathbf{H}}$ with eigenvalue $\lambda_0 = \sum_{i=0}^{\mu} h_i$.

(b) Show that row 2 of $\mathbf{Q}$ is an eigenvector of $\tilde{\mathbf{H}}$ with eigenvalue $\lambda_1 = \sum_{i=0}^{\mu} h_i W_N^i$.

(c) Argue by induction that similar relations hold for all rows of $\mathbf{Q}$.

10. Show that appending the all-zero prefix to an OFDM symbol and then adding in the tail of the received sequence, as shown in Figure 12.8, results in the same received sequence as with a cyclic prefix.

11. Consider a discrete-time FIR channel with $h[n] = .7\delta[n] + .5\delta[n-1] + .3\delta[n-3]$. Consider an OFDM system with $N = 8$ subchannels.

(a) Find the matrix $\mathbf{H}$ corresponding to the matrix representation of $\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{v}$ given in (12.25).

(b) Find the circulant convolution matrix $\mathbf{H}$ corresponding to the matrix representation in (12.27), as well as its eigenvalue decomposition $\mathbf{H} = \mathbf{M}\mathbf{A}\mathbf{M}^H$.

(c) What are the flat fading channel gains associated with each subchannel in the representation of part (b)?

(d) Consider now a point to point link with 2 Tx antennas and 2 Rx antennas. The discrete time FIR channels are as follows:

- From Tx 1 to Rx 1: $h[n] = 0.7\delta[n] + 0.5\delta[n-1] + 0.3\delta[n-3]$
- From Tx 2 to Rx 1: $h[n] = 0.1\delta[n] - 0.7\delta[n-1] + 0.6\delta[n-2]$
- From Tx 1 to Rx 2: $h[n] = 0.8\delta[n] - 0.2\delta[n-1]$
- From Tx 2 to Rx 2: $h[n] = -0.8\delta[n] - 0.1\delta[n-1]$

What are the channel matrices from the transmitters to the receivers (refer to (12.25)) ?

(e) How are the eigenvectors of the circulant forms of these matrices related to the $\mathbf{M}$ above? (*Note that you need not compute the eigendecomposition to answer this. Another hint: the eigenvectors of a general circulant matrix are the columns of a DFT matrix.*)

12. Consider a five-tap discrete-time channel

$$H(z) = 1 + .6z^{-1} + .7z^{-2} + .3z^{-3} + .2z^{-4}.$$

Assume this channel model characterizes the maximum delay spread of the channel, and assume a vector coded system is used over the channel with $N = 256$ carriers.

(a) What value of μ is needed for the prefix to eliminate ISI between vector codewords? What is the overhead associated with this μ?

(b) Find the system matrix representation (12.25) and the singular values of the associated channel matrix $\mathbf{H}$.

(c) Find the transmit precoding and receiver shaping matrices, $\mathbf{V}$ and $\mathbf{U}^H$, required to orthogonalize the subchannels.

13. Find the PAR of a raised cosine pulse with $\beta = 0, .5, 1$. Which pulse shape has the lowest PAR? Is this pulse shape more or less sensitive to timing errors?

14. Find the constant C_0 associated with intercarrier interference in (12.55).

15. Suppose the four subchannels in 802.11a used for pilot estimation could be used for data transmission by taking advantage of blind estimation techniques. Assuming the same modulation and coding formats are available, what maximum and minimum data rates could be achieved by including these extra subchannels?

16. Find the data rate of an 802.11a system assuming half the available 48 subchannels use BPSK with a rate-1/2 channel code and the others use 64-QAM with a rate-3/4 channel code.

17. For an OTFS system, show that the received symbols in the delay-Doppler domain $y[k, l]$ satisfy

$$y[k,l] = \frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[m,n] C_w \left(\frac{k-m}{MT}, \frac{l-n}{N\Delta\rho} \right) \tag{12.69}$$

where the windowed channel scattering function $C_w(\rho, \tau)$ is given by

$$C_w(\tau, \rho) = \iint c(x, y) w(\tau - x, \rho - y) e^{-j2\pi xy} dy dx. \tag{12.70}$$

Also show that for $C(\tau, \rho) = \delta(\tau)\delta(\rho)$, this reduces to $y[k, l] = c(0, 0)x[k, l]$ for any transmit and receive windowing functions.

Bibliography

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag*., pp. 5–14, May 1990.

[2] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag*., pp. 100–9, February 1995.

[3] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Sel. Areas Commun*., pp. 895–908, August 1991.

[4] J. M. Cioffi, "A multicarrier primer," Stanford University/Amati TIEI contribution, I1E1.4/91–157, November 1991.

[5] A. R. S. Bahai, B. R. Saltzberg, and M. Ergen, *Multi-Carrier Digital Communications-Theory and Applications of OFDM*, 2nd ed., Springer-Verlag, New York, 2004.

[6] IEEE 802.11a-l999: *High-Speed Physical Layer in the 5 GHz Band*, 1999.

[7] IEEE 802.11g-2003: *Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, 2003.

[8] IEEE 802.16a-200l: *IEEE Recommended Practice for Local and Metropolitan Area Networks*, 2001.

[9] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang, "IEEE Standard 802.16: A technical overview of the WirelessMAN 326 air interface for broadband wireless access," *IEEE Commun. Mag*., pp. 98–107, June 2002.

[10] M. Corson, R. Laroia, A. O'Neill, V. Park, and G. Tsirtsis, "A new paradigm for IP-based cellular networks," *IT Professional* pp. 20–9, November/December 2001.

[11] W. Lu, "4G mobile research in Asia," *IEEE Commun. Mag*., pp. 104–6, March 2003.

[12] T. S. Rappaport, A. Annamalai, R. M. Buehrer, and W. H. Tranter, "Wireless communications: Past events and a future perspective," *IEEE Commun. Mag*., pp. 148–61, May 2002.

[13] J. M. Cioffi. *Digital Communications*, chap. 4: *Multichannel Modulation*, unpublished course notes, available at ⟨http://www.stanford.edu/class/ee379c/⟩.

[14] S. Kasturia, J. Aslanis, and J. Cioffi, "Vector coding for partial response channels," *IEEE Trans. Inform. Theory*, pp. 741–62, July 1990.

[15] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Inform. Theory*, pp. 665–75, July 1985.

[16] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun*., pp. 773–5, February–April 1995.

[17] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Trans. Commun*., pp. 380–94, March 2004.

[18] R. K. Jurgen, "Broadcasting with digital audio," *IEEE Spectrum*, pp. 52–9, March 1996.

[19] S. Kaider, "Performance of multi-carrier CDM and COFDM in fading channels," *Proc. IEEE Globecom Conf*., pp. 847–51, December 1999.

[20] P. Hoeher, S. Kaiser, and P. Robertson, " Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," *Proc. IEEE Internat. Conf. Acous., Speech, Signal Proc*., pp. 1845–8, April 1997.

[21] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. I: Unification and optimal designs," *IEEE Trans. Signal Proc*., pp. 1988–2006, July 1999.

[22] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," IEEE J. Select. Areas Commun., Vol 16, No. 8, pp. 1579–1589, Aug. 1998.

[23] S. Das and P. Schniter, "Max-SINR ISI/ICI-shaping multicarrier communication over the doubly dispersive channel," *IEEE Trans. Signl. Proc.* Vol. 55, No.12, pp. 5782-5795, Dec. 2007.

[24] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. Commun.*, Vol. 51, No. 7 pp. 1111-1122, July 2003.

[25] Z. Zhao, M. Schellmann, X. Gong, Q. Wang, R. Böhnke, and Y. Guo, "Pulse shaping design for OFDM systems." *EURASIP Journal on Wireless Communications and Networking,* 2017, 1-25.

[26] H. Blcskei, P. Duhamel, and R. Hleiss, "Design of pulse shaping OFDM/OQAM systems for high data-rate transmission over wireless channels," *Proc. 1999 IEEE International Conference on Communications*, Vancouver, BC, Canada, 1999.

[27] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. II: Blind channel estimation, synchronization, and direct equalization," *IEEE Trans. Signal Proc*., pp. 2007–22, July 1999.

[28] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1999.

[29] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results," *IEEE Commun. Mag*., pp. 143–9, September 2002.

[30] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *Bell System Tech. J*., pp. 745–78, May/June 1974.

[31] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, pp. 271–94, February 2004.

[32] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun*., pp. 282–9, February 2001.

[33] D. J. G. Mestdagh and P. M. P. Spruyt, "A method to reduce the probability of clipping in DMT-based transceivers," *IEEE Trans. Commun*., pp. 1234–8, October 1996.

[34] K. G. Paterson and V. Tarokh, "On the existence and construction of good codes with low peak-to-average power ratios," *IEEE Trans. Inform. Theory*, pp. 1974–87, September 2000.

[35] J. Tellado, *Multicarrier Modulation with Low PAR: Applications to DSL and Wireless*, Kluwer, Boston, 2000.

[36] C. Muschallik, "Improving an OFDM reception using an adaptive Nyquist windowing," *IEEE Trans. Consumer Elec*., pp. 259–69, August 1996.

[37] A. Redfern, "Receiver window design for multicarrier communication systems," *IEEE J. Sel. Areas Commun*., pp. 1029–36, June 2002.

[38] T. Wang, J. G. Proakis, E. Masry, and J. Zeidler, "Performance degradation of OFDM systems due to Doppler spreading", *IEEE Trans. Commun.*, Vol. 5, No. 6, pp. 1422–1432, June 2006.

[39] T. Dean, M. Chowdhury, N. Grimwood, A. Goldsmith, "Rethinking modulation and detection for high doppler channels," *IEEE Transactions on Wireless Communications*, Jan. 2020.

[40] A. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, Vol. 47, No. 1, pp. 123–132, Jan. 1999.

[41] K. Liu, T. Kadous, and A.M. Sayeed, "Orthogonal time-frequency signaling over doubly dispersive channels, *IEEE Trans. Inform. Theory*, Vol. 50, No. 11, pp. 2583–2603, Nov. 2004.

[42] X. Ma and G. B. Giannakis, "Maximum-diversity transmissions Over doubly selective wireless channels," *IEEE Trans. Inform. Theory*, Vol. 49, No. 7, pp. 1832–1840, July 2003.

[43] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 136-144, August 2021.

[44] Y. Hong, T. Tharaj, and E. Viterbo. *Delay-Doppler Communications: Principles and Applications*. Academic Press, 2022.

[45] Z. Wei, S. Li, W. Yuan, R. Schober, and G. Caire, "Orthogonal time frequency space modulation—Part I: Fundamentals and challenges ahead; Part II: Transceiver designs; Part III: ISAC and potential applications" *IEEE Communications Letters*, vol. 27, no. 1, pp. 4-13, 2022.

[46] R. Hadani et al., "Orthogonal time frequency space modulation," *Proc. 2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017.

[47] P. Raviteja, K. T. Phan, Y. Hong and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6501-6515, Oct. 2018.

[48] S. K. Mohammed, R. Hadani, A. Chockalingam and R. Calderbank, "OTFS—predictability in the delay-Doppler domain and its value to communication and radar sensing," *IEEE BITS the Information Theory Magazine*, vol. 3, no. 2, pp. 7-31, June 2023.

[49] S. K. Mohammed, R. Hadani, and A. Chockalingam, *OTFS modulation: Theory and applications*, Wiley-IEEE Press, 2024.

Chapter 13

Spread Spectrum

Although bandwidth is a valuable commodity in wireless systems, *increasing* the transmit signal bandwidth can sometimes improve performance. Spread spectrum is a technique that increases signal bandwidth beyond the minimum necessary for data communication. There are many reasons for doing this. Spread-spectrum techniques can hide a signal below the noise floor, making it difficult to detect. Spread spectrum also mitigates the performance degradation due to intersymbol and narrowband interference. In conjunction with a RAKE receiver, spread spectrum can provide coherent combining of different multipath components. The wide bandwidth of spread-spectrum signals is useful for location and timing acquisition. Finally, spread spectrum is a *multiple access* technique; it allows multiple users to share the same signal bandwidth since spread signals can be superimposed on top of each other. The use of spread spectrum for multiple access will be discussed in Sections 13.4-13.5 and its performance relative to other multiple access techniques will be discussed in Chapter 14.

Spread spectrum first achieved widespread use in military applications because of its inherent property of hiding the spread signal below the noise floor during transmission, its resistance to narrowband jamming and interference, and its low probability of detection and interception. For commercial applications, cordless phones typically employ spread spectrum due to its narrowband interference resistance properties. Second and third generation cellular systems as well as second-generation wireless LANs adopted spread spectrum for its ISI rejection and bandwidth-sharing capabilities. Later generations of both cellular and wireless LANs switched to OFDM-based multiple access techniques, which are discussed in the next chapter.

13.1 Spread-Spectrum Principles

Spread spectrum is a modulation method applied to digitally modulated signals that increases the transmit signal bandwidth to a value much larger than is needed to transmit the underlying information bits. There are many signaling techniques that increase the transmit bandwidth above the minimum required for data transmission – for example, coding and frequency modulation. However, these techniques do not fall in the category of spread spectrum. The following three properties are needed for a signal to be spread-spectrum modulated [1].

1. The signal occupies a bandwidth much larger than is needed for the information signal.
2. The spread-spectrum modulation is done using a *spreading code*, which is independent of the data in the signal.
3. Despreading at the receiver is done by correlating the received signal with a synchronized copy of the spreading code.

In order to make these notions precise, we return to the signal space representation of Section 5.1 to investigate embedding an information signal of bandwidth B within a much larger transmit signal bandwidth B_s than is

needed. From (5.3), a set of linearly independent signals $s_i(t)$, $i = 1, \ldots, M$, of bandwidth B and time duration T can be written using a basis function representation as

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t), \qquad 0 \le t < T, \tag{13.1}$$

where the basis functions $\phi_j(t)$ are orthonormal and span an N-dimensional space. One of these signals is transmitted every T seconds to convey $\log_2 M/T$ bits per second. As discussed in Section 5.1.2, the minimum number of basis functions needed to represent these signals is approximately $2BT$. Since the $\{s_i(t)\}_{i=1}^{M}$ are linearly independent, this implies $M \approx 2BT$. To embed these signals into a higher-dimensional space, we choose $N \gg M$. The receiver uses an M-branch structure where the ith branch correlates the received signal with $s_i(t)$. The receiver outputs the signal corresponding to the branch with the maximum correlator output.

Suppose we generate the signals $s_i(t)$ using *random sequences*, so that the sequence of coefficients s_{ij} are chosen based on a random sequence generation, where each coefficient has mean zero and variance E_s/N. Thus, the signals $s_i(t)$ will have their energies uniformly distributed over the signal space of dimension N. Consider an interference or jamming signal within this signal space. The interfering signal can be represented as

$$I(t) = \sum_{j=1}^{N} I_j\phi_j(t), \tag{13.2}$$

with total energy over $[0,\ T]$ given by

$$\int_0^T I^2(t)\,dt = \sum_{j=1}^{N} I_j^2 = E_J. \tag{13.3}$$

Suppose the signal $s_i(t)$ is transmitted. Neglecting noise, the received signal is the sum of the transmitted signal plus interference:

$$x(t) = s_i(t) + I(t). \tag{13.4}$$

The output of the correlator in the ith branch of the receiver is then

$$x_i = \int_0^T x(t)s_i(t)\,dt = \sum_{j=1}^{N}(s_{ij}^2 + I_j s_{ij}), \tag{13.5}$$

where the first term in this expression represents the signal and the second term represents the interference. It can be shown [1] that the signal-to-interference power ratio (SIR) of this signal is

$$\text{SIR} = \frac{E_s}{E_J} \cdot \frac{N}{M}. \tag{13.6}$$

This result is independent of the distribution of the interferer's energy over the N-dimensional signal space. In other words, by spreading the interference power over a larger dimension N than the required signaling dimension M, the SIR is increased by $G = N/M$, where G is called the *processing gain* or *spreading factor*. In practice, spread-spectrum systems have processing gains on the order of 10–1000. Since $N \approx 2B_sT$ and $M \approx 2BT$, we have $G \approx B_s/B$, the ratio of the spread signal bandwidth to the information signal bandwidth. Processing gain is often defined as this bandwidth ratio or something similar, but its underlying meaning is generally related to the performance improvement of a spread-spectrum system relative to a non-spread system in the presence of interference [2, Chap. 2.1]. Note that block and convolution coding are also techniques that improve performance in

the presence of noise or interference by increasing signal bandwidth. An interesting trade-off arises as to whether, given a specific spreading bandwidth, it is more beneficial to use coding or spread spectrum. The answer depends on the specifics of the system design [3].

Spread spectrum is typically implemented in one of two forms: *direct sequence* (DS) or *frequency hopping* (FH). In direct-sequence spread-spectrum (DSSS) modulation, the modulated data signal $s(t)$ is multiplied by a wideband *spreading signal* or *code* $s_c(t)$, where $s_c(t)$ is constant over a time duration T_c and has amplitude equal to 1 or –1. The spreading code bits are usually referred to as *chips*, T_c is called the *chip time*, and $1/T_c$ is called the *chip rate*. The bandwidth $B_c \approx 1/T_c$ of $s_c(t)$ is roughly $B_c/B \approx T_s/T_c$ times larger than the bandwidth B of the modulated signal $s(t)$, and the number of chips per bit, T_s/T_c, is an integer approximately equal to G, the processing gain of the system. Multiplying the modulated signal by the spreading signal results in the convolution of these two signals in the frequency domain. Thus, the transmitted signal $s(t)s_c(t)$ has frequency response $S(f) * S_c(f)$, which has a bandwidth of roughly $B_c + B$. The multiplication of a spreading signal $s_c(t)$ with a modulated data signal $s(t)$ over one symbol time T_s is illustrated in Figure 13.1.

For an AWGN (additive white Gaussian noise) channel, the received spread signal is $s(t)s_c(t) + n(t)$ for $n(t)$ the channel noise. If the receiver multiplies this signal by a synchronized replica of the spreading signal, the result is $s(t)s_c^2(t) + n(t)s_c(t)$. Since $s_c(t) = \pm 1$, we have $s_c^2(t) = 1$. Moreover, $n'(t) = n(t)s_c(t)$ has approximately the same white Gaussian statistics as $n(t)$ if $s_c(t)$ is zero mean and sufficiently wideband (i.e., if its autocorrelation approximates a delta function). Thus, the received signal is $s(t)s_c^2(t) + n(t)s_c(t) = s(t) + n'(t)$, which indicates that spreading and despreading have no impact on signals transmitted over AWGN channels. However, spreading and despreading have tremendous benefits when the channel introduces narrowband interference or ISI.

We now illustrate the narrowband interference and multipath rejection properties of direct-sequence spread spectrum (DSSS) in the frequency domain: more details will be given in later sections. We first consider narrowband interference rejection, as shown in Figure 13.2. Neglecting noise, we see that the receiver input consists of the spread modulated signal $S(f) * S_c(f)$ and the narrowband interference $I(f)$. The despreading in the receiver recovers the data signal $S(f)$. However, the interference signal $I(t)$ is multiplied by the spreading signal $s_c(t)$, resulting in their convolution $I(f) * S_c(f)$ in the frequency domain. Thus, receiver despreading has the effect of distributing the interference power over the bandwidth of the spreading code. The demodulation of the modulated signal $s(t)$ effectively acts as a lowpass filter, removing most of the energy of the spread interference, which reduces its power by the processing gain $G \approx B_c/B$.

Figure 13.3 illustrates ISI rejection, which is based on a similar premise. Suppose the spread signal $s(t)s_c(t)$ is transmitted through a two-ray channel with impulse response $h(t) = \alpha\delta(t) + \beta\delta(t - \tau)$. Then $H(f) = \alpha + \beta e^{-j2\pi f\tau}$, resulting in a receiver input in the absence of noise equal to $H(f)[S(f) * S_c(f)]$ in the frequency domain or $[s(t)s_c(t)] * h(t) = \alpha s(t)s_c(t) + \beta s(t - \tau)s_c(t - \tau)$ in the time domain. Suppose that the receiver despreading process multiplies this signal by a copy of $s_c(t)$ synchronized to the first path of this two-ray model. This results in the time-domain signal $\alpha s(t)s_c^2(t) + \beta s(t-\tau)s_c(t-\tau)s_c(t)$. Since the second multipath component $\beta s'(t) = \beta s(t-\tau)s_c(t-\tau)s_c(t)$ includes the product of asynchronized copies of $s_c(t)$, it remains spread out over the spreading code bandwidth, and the demodulation process will remove most of its energy. More precisely, as described in Section 13.2, the demodulation process effectively attenuates the multipath component by the autocorrelation $\rho_c(\tau)$ of the spreading code at delay τ. This autocorrelation can be quite small when $\tau > T_c$, on the order of $1/G \approx T_c/T_s$, resulting in significant mitigation of the ISI when the modulated signal is spread over a wide bandwidth. Since the spreading code autocorrelation determines the ISI rejection of the spread-spectrum system, it is important to use spreading codes with good autocorrelation properties, as discussed in the next section.

The basic premise of frequency-hopping spread spectrum (FHSS) is to hop the modulated data signal over a wide bandwidth by changing its carrier frequency according to the value of a spreading code $s_c(t)$.[1] This process

[1]The concept of frequency hopping was invented during World War II by film star Hedy Lamarr and composer George Antheil. Their patent for a "secret communications system" used a chip sequence generated by a player piano roll to hop between 88 frequencies. The design was intended to make radio-guided torpedoes hard to detect jam.

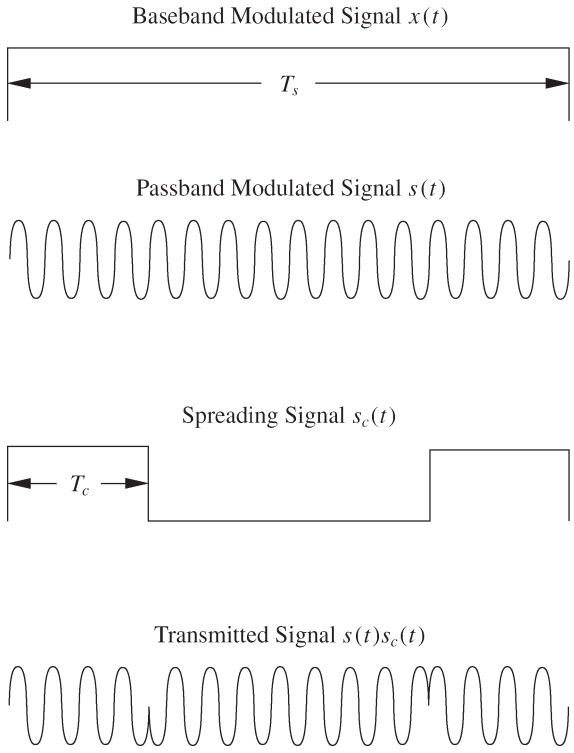


Figure 13.1: Spreading signal multiplication.

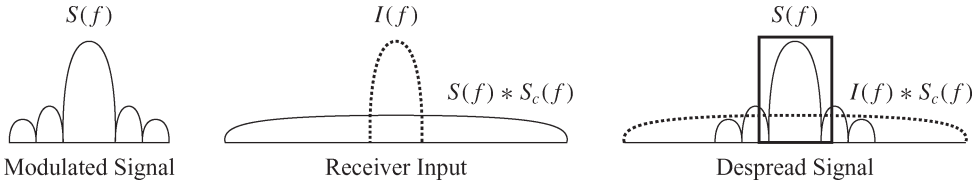


Figure 13.2: Narrowband interference rejection in DSSS.

is illustrated in Figure 13.4. The chip time T_c of $s_c(t)$ dictates the time between hops – that is, the time duration over which the modulated data signal is centered at a given carrier frequency f_i before hopping to a new carrier frequency. The hop time can exceed a symbol time, $T_c = kT_s$ for some integer k, which is called slow frequency hopping (SFH); or the carrier can be changed multiple times per symbol, $T_c = T_s/k$ for some integer k, which is called fast frequency hopping (FFH). In FFH there is frequency diversity on every symbol, which protects each symbol against narrowband interference and spectral nulls due to frequency-selective fading. The bandwidth of the FH system is approximately equal to NB, where N is the number of carrier frequencies available for hopping and B is the bandwidth of the data signal. The signal is generated using a frequency synthesizer that determines the modulating carrier frequency from the chip sequence, typically using a form of FM modulation such as CPFSK. In the receiver, the signal is demodulated using a similar frequency synthesizer, synchronized to the chip sequence $s_c(t)$, that generates the sequence of carrier frequencies from this chip sequence for downconversion. As with DS, FH has no impact on performance in an AWGN channel. However, it does mitigate the effects of narrowband interference and multipath.

Consider a narrowband interferer of bandwidth B at a carrier frequency f_i corresponding to one of the carriers used by the FH system. The interferer and FH signal occupy the same bandwidth only when carrier f_i is generated

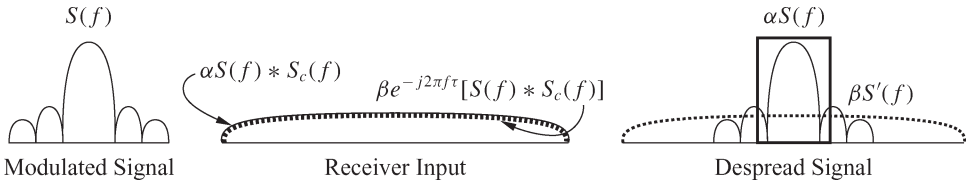


Figure 13.3: ISI rejection in DSSS.

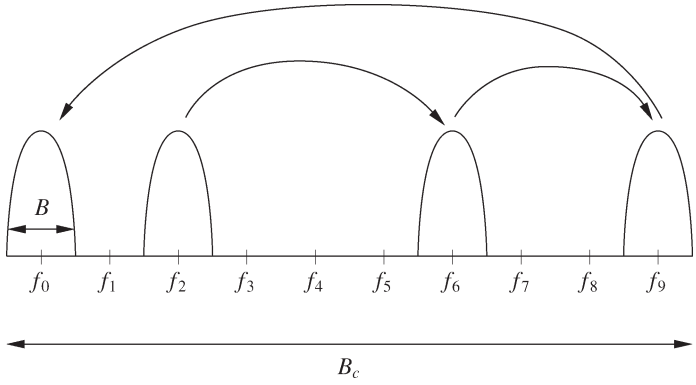


Figure 13.4: Frequency hopping.

by the hop sequence. If the hop sequence spends an equal amount of time at each of the carrier frequencies, then interference occurs a fraction $1/N$ of the time, and thus the interference power is reduced by roughly $1/N$. However, the nature of the interference reduction is different in FH versus DS systems. In particular, DS results in a reduced-power interference all the time, whereas FH has a full power interferer a fraction of the time. In FFH systems the interference affects only a fraction of a symbol time, so coding may not be required to compensate for this interference. In SFH systems the interference affects many symbols, so typically coding with interleaving is needed to avoid many simultaneous errors in a single codeword. Frequency hopping is commonly used in military systems, where the interferers are assumed to be malicious jammers attempting to disrupt communications.

We now investigate the impact of multipath on an FH system. For simplicity, we consider a two-ray channel that introduces a multipath component with delay τ. Suppose the receiver synchronizes to the hop sequence associated with the line-of-sight signal path. Then the LOS path is demodulated at the desired carrier frequency. However, the multipath component arrives at the receiver with a delay τ. If $\tau > T_c$ then the receiver will have hopped to a new carrier frequency $f_j \neq f_i$ for downconversion when the multipath component, centered at carrier frequency f_i, arrives at the receiver. Since the multipath occupies a different frequency band than the LOS signal component being demodulated, it causes negligible interference to the demodulated signal. Thus, the demodulated signal does not exhibit either flat or frequency-selective fading for $\tau > T_c$. If $\tau < T_c$ then the impact of multipath depends on the bandwidth B of the modulated data signal as well as on the hop rate. First consider an FFH system where $T_c \ll T_s$. Since we also assume $\tau < T_c$, it follows that $\tau < T_c \ll T_s$. Since all the multipath arrives within a symbol time, the multipath introduces a complex amplitude gain and the signal experiences flat fading. Now consider an SFH system where $T_c \gg T_s$. Since we also assume $\tau < T_c$, all the multipath will arrive while the signal is at the same carrier frequency, so the impact of multipath is the same as if there were no frequency hopping: for $B < 1/\tau$ the signal experiences flat fading, and for $B > 1/\tau$ the signal experiences frequency-selective fading. The fading channel also varies slowly over time, since the equivalent lowpass channel is a function of the carrier frequency and thus changes whenever the carrier hops to a new frequency. In summary, frequency hopping removes the impact of multipath on demodulation of the LOS component whenever $\tau > T_c$.

For $\tau < T_c$, an FFH system will exhibit flat fading, whereas an SFH system will exhibit slowly varying flat fading for $B < 1/\tau$ and slowly varying frequency-selective fading for $B > 1/\tau$. The performance analysis under time-varying flat or frequency-selective fading is the same as for systems without hopping, as given in Sections 6.3 and 6.5, respectively.

In addition to their interference and ISI rejection capabilities, both DSSS and FHSS provide a mechanism for multiple access, allowing many users to simultaneously share the spread bandwidth with minimal interference between users. In these multiuser systems, the interference between users is determined by the cross-correlation of their spreading codes. Spreading code designs typically have either good autocorrelation properties to mitigate ISI or good cross-correlation properties to mitigate multiuser interference. However, there is usually a trade-off between optimizing the two features. Thus, the best choice of code design depends on the number of users in the system and the severity of the multipath and interference. Trade-offs between frequency hopping and direct sequence in multiuser systems are discussed in Section 13.5. Frequency hopping is also used in cellular systems to average out interference from other cells.

Example 13.1: Consider an SFH system with hop time $T_c = 10$ μs and symbol time $T_s = 1$ μs. If the FH signal is transmitted over a multipath channel, for approximately what range of multipath delay spreads will the received despread signal exhibit frequency-selective fading?

Solution: Based on the two-ray model analysis, the signal exhibits fading – flat or frequency-selective – only when the delay spread $\tau < T_c = 10$ μs. Moreover, for frequency-selective fading we require $B \approx 1/T_s = 10^6 > 1/\tau$; that is, we require $\tau > 10^{-6} = 1$ μs. So the despread signal will exhibit frequency-selective fading for delay spreads ranging from approximately 1 μs to 10 μs.

13.2 Direct-Sequence Spread Spectrum (DSSS)

13.2.1 DSSS System Model

An end-to-end direct-sequence spread-spectrum system is illustrated in Figure 13.5. The multiplication by $s_c(t)$ and the carrier $\cos(2\pi f_c t)$ could be done in opposite order as well: downconverting prior to despreading allows the code synchronization and despreading to be done digitally, but it complicates carrier phase tracking because this must be done relative to the wideband spread signal.[2] For simplicity we illustrate the receiver only for in-phase signaling; a similar structure is used for the quadrature signal component. The data symbols s_l are first linearly modulated to form the baseband modulated signal $x(t) = \sum_l s_l g(t - lT_s)$, where $g(t)$ is the modulator shaping pulse, T_s the symbol time, and s_l the symbol transmitted over the lth symbol time. Linear modulation is used because DSSS is a form of phase modulation and therefore works best in conjunction with a linearly modulated data signal. The modulated signal is then multiplied by the spreading code $s_c(t)$ with chip time T_c, after which it is upconverted through multiplication by the carrier $\cos(2\pi f_c t)$. The spread signal passes through the channel

[2] A system that performs spreading and despreading on the bandpass modulated signal works as follows. The transmitter consists of a standard narrowband modulator that generates a passband modulated signal, followed by spreading. The receiver consists of despreading, followed by a standard narrowband demodulator. This order of operations makes it straightforward to design a spread-spectrum system using existing narrowband modulators and demodulators, and operations such as carrier phase recovery are not affected by spreading. However, spread-spectrum systems today do as much of the signal processing as possible in the digital domain. Thus, spread-spectrum systems typically modulate the data symbols and multiply by the spreading code at baseband using digital signal processing, followed by A/D conversion and analog upconversion to the carrier frequency. In this case all functions prior to the carrier multiplication in Figure 13.5 are done digitally, with an A/D converter following the multiplication with $s_c(t)$. However, the carrier recovery loop for this system is more difficult to design because it operates on the spread signal. In particular, any nonlinear operation (e.g. squaring) that is used to remove either the data or the spreading sequence in carrier phase recovery can be seriously degraded by the noise associated with the spread signal.

$h(t)$, which also introduces narrowband interference $I(t)$ and zero-mean AWGN $n(t)$ with power spectral density $N_0/2$.

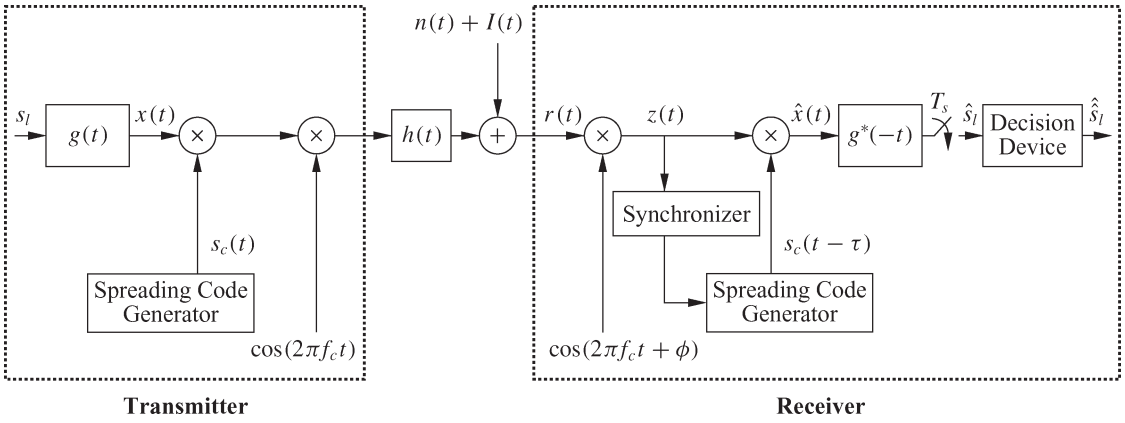


Figure 13.5: DSSS system model.

Assume the channel introduces several multipath components: $h(t) = \alpha_0\delta(t-\tau_0) + \alpha_1\delta(t-\tau_1) + \cdots$. The received signal is first downconverted to baseband assuming perfect carrier recovery, so that the carrier $\cos(2\pi f_c t + \phi)$ in the receiver has phase ϕ matched to the carrier phase of the incoming signal (e.g., if $h(t) = \delta(t)$ then $\phi = 0$). The carrier recovery loop is typically assumed to lock to the carrier associated with the LOS (or minimum delay) multipath component. The synchronizer then uses the resulting baseband signal $z(t)$ to align the delay τ of the receiver's spreading code generator with one of the multipath component delays τ_i. The spreading code generator then outputs the spreading code $s_c(t-\tau)$, where $\tau = \tau_i$ if the synchronizer is perfectly aligned with the delay associated with the ith multipath component. Ideally the synchronizer would lock to the multipath component with the largest amplitude. However, in practice this requires a complex search procedure, so instead the synchronizer typically locks to the first component it finds with an amplitude above a given threshold. This synchronization procedure can be quite complex, especially for channels with severe ISI or interference, and synchronization circuitry can make up a large part of any spread-spectrum receiver. Synchronization is discussed in more detail in Section 13.2.3.

The multipath component at delay τ is despread via multiplication by the spreading code $s_c(t-\tau)$. The other multipath components are not despread, and most of their energy is removed, as we shortly show. After despreading, the baseband signal $\hat{x}(t)$ passes through a matched filter and decision device. Thus, there are three stages in the receiver demodulation for direct-sequence spread spectrum: downconversion, despreading, and baseband demodulation. This demodulator is also called the single-user *matched filter* detector for DSSS. We now examine the three stages of this detector in more detail.

For simplicity, assume rectangular pulses are used in the modulation ($g(t) = \sqrt{2/T_s},\ 0 \le t \le T_s$). The matched filter $g^*(-t)$ then simply multiplies $\hat{x}(t)$ by $\sqrt{2/T_s}$ and integrates from zero to T_s to obtain the estimate of the transmitted symbol. Since perfect carrier recovery is assumed, the carrier phase offset ϕ in the receiver matches that of the incoming signal. We also assume perfect synchronization in the receiver. The multipath and interference rejection occurs in the data demodulation process. Specifically, the input to the matched filter is given by

$$\begin{aligned}\hat{x}(t) =& ([x(t)s_c(t)\cos(2\pi f_c t)] * h(t))s_c(t-\tau)\cos(2\pi f_c t + \phi) \\ &+ n(t)s_c(t-\tau)\cos(2\pi f_c t + \phi) + I(t)s_c(t-\tau)\cos(2\pi f_c t + \phi). \end{aligned} \tag{13.7}$$

Without multipath, $h(t) = \delta(t)$, $\phi = 0$, and the receiver ideally synchronizes with $\tau = 0$. Then the spreading/despreading process has no impact on the baseband signal $x(t)$. Specifically, the spreading code has amplitude ± 1, so multiplying $s_c(t)$ by a synchronized copy of itself yields $s_c^2(t) = 1$ for all t. Then, in the absence of multipath and interference (i.e., for $h(t) = \delta(t)$ and $I(t) = 0$),

$$\begin{aligned}\hat{x}(t) &= x(t)s_c^2(t)\cos^2(2\pi f_c t) + n(t)s_c(t)\cos(2\pi f_c t) \\ &= x(t)\cos^2(2\pi f_c t) + n(t)s_c(t)\cos(2\pi f_c t),\end{aligned} \tag{13.8}$$

since $s_c^2(t) = 1$. If $s_c(t)$ is sufficiently wideband then $n(t)s_c(t)$ has approximately the same statistics as $n(t)$: it is a zero-mean AWGN random process with PSD $N_0/2$. The matched filter output over a symbol time for $g(t) = \sqrt{2/T_s},\ 0 \le t \le T_s$, will thus be

$$\begin{aligned}\hat{s}_l &= \int_0^{T_s} \hat{x}(t) * g^*(-t)\,dt \\ &= \sqrt{\frac{2}{T_s}}\int_0^{T_s} x(t)\cos^2(2\pi f_c t)\,dt + \sqrt{\frac{2}{T_s}}\int_0^{T_s} n(t)s_c(t)\cos(2\pi f_c t)\,dt \\ &= \frac{2}{T_s}\int_0^{T_s} s_l\cos^2(2\pi f_c t)\,dt + \sqrt{\frac{2}{T_s}}\int_0^{T_s} n(t)s_c(t)\cos(2\pi f_c t)\,dt \\ &\approx s_l + n_l,\end{aligned} \tag{13.9}$$

where s_l and n_l correspond to the data and noise output of a standard demodulator without spreading or despreading and where the approximation assumes $f_c \gg 1/T_s$.

We now consider the interference signal $I(t)$ at the carrier frequency f_c, which can be modeled as $I(t) = I'(t)\cos(2\pi f_c t)$ for some narrowband baseband signal $I'(t)$. We again assume $h(t) = \delta(t)$. Multiplication by the spreading signal perfectly synchronized to the incoming signal yields

$$\hat{x}(t) = x(t)\cos^2(2\pi f_c t) + n(t)s_c(t)\cos(2\pi f_c t) + I'(t)s_c(t)\cos^2(2\pi f_c t), \tag{13.10}$$

where $n(t)s_c(t)$ is assumed to be a zero-mean AWGN process. The demodulator output is then given by

$$\begin{aligned}\hat{s}_l &= \frac{2}{T_s}\int_0^{T_s} s_l s_c^2(t)\cos^2(2\pi f_c t)\,dt + \sqrt{\frac{2}{T_s}}\int_0^{T_s} n(t)s_c(t)\cos(2\pi f_c t)\,dt \\ &\quad + \sqrt{\frac{2}{T_s}}\int_0^{T_s} I'(t)s_c(t)\cos^2(2\pi f_c t)\,dt \\ &\approx s_l + n_l + I_l,\end{aligned} \tag{13.11}$$

where s_l and n_l correspond to the data and noise output of a standard demodulator without spreading or despreading and where the approximation assumes $f_c \gg 1/T_s$. The narrowband interference rejection can be seen from the last term of (13.11). In particular, the spread interference $I'(t)s_c(t)$ is a wideband signal with bandwidth of roughly $1/T_c$, and the integration acts as a lowpass filter with bandwidth of roughly $1/T_s \ll 1/T_c$, thereby removing most of the interference power.

Let us now consider ISI rejection. Assume a multipath channel with one delayed component: $h(t) = \alpha_0\delta(t) + \alpha_1\delta(t - \tau_1)$. For simplicity, assume $\tau_1 = kT_s$ is an integer multiple of the symbol time. Suppose that the first multipath component is stronger than the second, $\alpha_0 > \alpha_1$, and that the receiver locks to the first component

($\phi = 0$ and $\tau = 0$ in Figure 13.5). Then, in the absence of narrowband interference ($I(t) = 0$), after despreading we have

$$\begin{aligned}\hat{x}(t) =& \alpha_0 x(t) s_c^2(t) \cos^2(2\pi f_c t) \\ &+ \alpha_1 x(t-\tau_1) s_c(t-\tau_1) \cos(2\pi f_c (t-\tau_1)) s_c(t) \cos(2\pi f_c t) \\ &+ n(t) s_c(t) \cos(2\pi f_c t). \end{aligned} \tag{13.12}$$

Since $\tau_1 = kT_s$, the ISI just corresponds to the signal transmission of the $(l-k)$th symbol; that is, $x(t-\tau_1) = x(t-kT_s) = s_{l-k} g(t-(l-k)T_s)$. The demodulator output over the lth symbol time is then given by

$$\begin{aligned}\hat{s}_l =& \frac{2}{T_s} \int_0^{T_s} \alpha_0 s_l \cos^2(2\pi f_c t)\, dt \\ &+ \frac{2}{T_s} \int_0^{T_S} \alpha_1 s_{l-k} s_c(t) s_c(t-\tau_1) \cos(2\pi f_c t) \cos(2\pi f_c (t-\tau_1))\, dt \\ &+ \sqrt{\frac{2}{T_s}} \int_0^{T_s} n(t) s_c(t) \cos(2\pi f_c t)\, dt \qquad (13.13) \\ \approx& \alpha_0 s_l + \alpha_1 s_{l-k} \cos(2\pi f_c \tau_1) \rho_c(\tau_1) + n_l; \qquad (13.14) \end{aligned}$$

here, as in the case of interference rejection, s_l and n_l correspond to the data symbol and noise output of a standard demodulator without spreading or despreading, and the approximation assumes $f_c \gg 1/T_s$. The middle term $\alpha_1 s_{l-k} \cos(2\pi f_c \tau_1) \rho_c(\tau_1)$ comes from the following integration:

$$\begin{aligned}&\frac{2}{T_s} \int_0^{T_s} s_c(t) s_c(t-\tau_1) \cos(2\pi f_c t) \cos(2\pi f_c (t-\tau_1))\, dt \\ &\quad = \frac{1}{T_s} \int_0^{T_s} s_c(t) s_c(t-\tau_1) (\cos(2\pi f_c \tau_1) + \cos(4\pi f_c t - 2\pi f_c \tau_1))\, dt \\ &\quad \approx \cos(2\pi f_c \tau_1) \frac{1}{T_s} \int_0^{T_s} s_c(t) s_c(t-\tau_1)\, dt \\ &\quad = \cos(2\pi f_c \tau_1) \rho_c(t,\ t-\tau_1), \end{aligned} \tag{13.15}$$

where the approximation is based on $f_c \gg T_s^{-1}$ and where

$$\rho_c(t,\ t-\tau_1) \triangleq \frac{1}{T_s} \int_0^{T_s} s_c(t) s_c(t-\tau_1)\, dt \tag{13.16}$$

is defined as the *autocorrelation* of the spreading code at time t and delay τ_1 over a symbol time.[3] More generally, the spreading code autocorrelation at time t (assumed to be an integer multiple of T_c) and delay τ, over a period $[0,\ T]$, is defined as

[3] Note that, if τ_1 is not an integer multiple of a symbol time, then the middle term in (13.14) becomes more complicated. In particular, assuming $g(t) = \sqrt{2/T_s}$, if $\tau_1 = (k+\kappa)T_s$ for $0 < \kappa < 1$ then $x(t-\tau_1) = \sqrt{2/T_s} s_{l-k-1}$ for $0 \le t \le \kappa T_s$ and $x(t-\tau_1) = \sqrt{2/T_s} s_{l-k}$ for $\kappa T_s \le t \le T_s$. Thus, the middle term of (13.14) becomes

$$\alpha_1 s_{l-k-1} \cos(2\pi f_c \tau_1) \frac{1}{T_s} \int_0^{\kappa T_s} s_c(t) s_c(t-\tau_1)\, dt + \alpha_1 s_{l-k} \cos(2\pi f_c \tau_1) \frac{1}{T_s} \int_{\kappa T_s}^{T_s} s_c(t) s_c(t-\tau_1)\, dt,$$

where each term is a function of the spreading code autocorrelation taken over a fraction of the symbol time.

$$\rho_c(t,\ t-\tau) \triangleq \frac{1}{T}\int_0^T s_c(t)s_c(t-\tau)\,dt = \frac{1}{N_T}\sum_{n=1}^{N_T} s_c(nT_c)s_c(nT_c-\tau), \tag{13.17}$$

where $N_T = T/T_c$ is the number of chips over duration T and where the second equality follows because $s_c(t)$ is constant over a chip time T_c. It can be shown that $\rho_c(t,\ t-\tau)$ is a symmetric function with maximum value at $\tau = 0$. Moreover, if $s_c(t)$ is periodic with period T, then the autocorrelation depends only on the time difference of the spreading codes; that is,

$$\frac{1}{T}\int_0^T s_c(t-\tau_0)s_c(t-\tau_1)\,dt = \rho_c(\tau_1-\tau_0), \tag{13.18}$$

so that $\rho_c(t,\ t-\tau) \triangleq \rho_c(\tau)$ in (13.17) depends only on τ.

By (13.15), if $T = T_s$ and $\rho_c(\tau) = \delta(\tau)$ then the despreading process removes all ISI. Unfortunately, it is not possible to have finite-length spreading codes with autocorrelation equal to a delta function. Hence there has been much work on designing spreading codes with autocorrelation over a symbol time that approximates a delta function. In the next section we discuss spreading codes for ISI rejection, including maximal linear codes, which have excellent autocorrelation properties to minimize ISI effects.

13.2.2 Spreading Codes for ISI Rejection: Random, Pseudorandom, and m-Sequences

Spreading codes are generated deterministically, often using a shift register with feedback logic to create a binary code sequence $\mathbf{b}$ of 1s and 0s. The binary sequence, also called a *chip sequence*, is used to amplitude modulate a square pulse train with pulses of duration T_c – with amplitude 1 for a 1-bit and amplitude -1 for a 0-bit – as shown in Figure 13.6. The resulting spreading code $s_c(t)$ is a sinc function in the frequency domain, corresponding to the Fourier transform of a square pulse. The shift register, consisting of n stages, has a cyclical output with a maximum period of $2^n - 1$. To avoid a spectral spike at DC or biasing the noise in despreading, the spreading code $s_c(t)$ should have no DC component, which requires that the bit sequence $\mathbf{b}$ have approximately the same number of 1s and 0s. It is also desirable for the number of consecutive 1s or 0s, called a *run*, to be small. Runs are undesirable because, if there is a run of k consecutive 1s or 0s, then the data signal over kT_c is just multiplied by a constant, which reduces the bandwidth spreading (and its advantages) by roughly a factor of k. Ideally the chip values change roughly every chip time, which leads to maximal spreading. Based on (13.15), we require spreading codes with $\rho_c(\tau) \approx \delta(\tau)$ in order to minimize ISI effects.

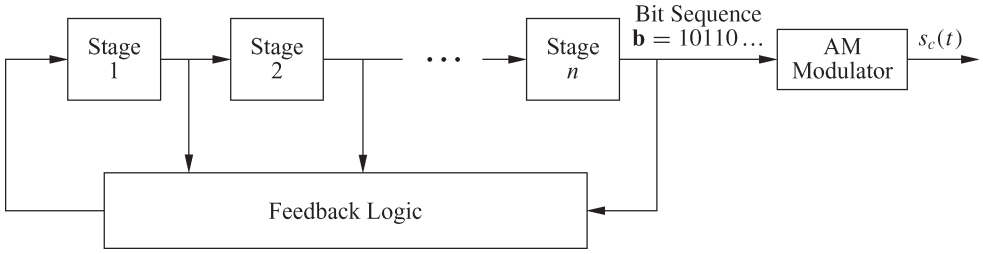


Figure 13.6: Generation of spreading codes.

Example 13.2: Find the baseband bandwidth of a spreading code $s_c(t)$ with chip time $T_c = 1\ \mu$s.
Solution: The spreading code $s_c(t)$ consists of a sequence of unit-amplitude square pulses of duration T_c modulated with ± 1. The Fourier transform of a unit-amplitude square pulse is $S(f) = T_c\ \text{sinc}\ (fT_c)$, with a main

lobe of bandwidth $2/T_c$. Thus, the null-to-null baseband bandwidth, defined as the minimum frequency where $S(f) = 0$, is $1/T_c$.

Although DSSS chip sequences must be generated deterministically, properties of random sequences are useful for gaining insight into deterministic sequence design. A random binary chip sequence consists of independent and identically distributed bit values with probability .5 for a l-bit or a 0-bit. A random sequence of length N can thus be generated, for example, by flipping a fair coin N times and setting the bit to a 1 for heads and a 0 for tails. Random sequences with length N asymptotically large have a number of the properties desired in spreading codes [4]. In particular, such sequences will have an equal number of 1s and 0s, called the *balanced property* of a code. Moreover, the run length in such sequences is generally short. In particular, for asymptotically large sequences, half of all runs are of length 1, a quarter are of length 2, and so forth, so that a fraction $1/2^r$ of all runs are of length r for r finite. This distribution on run length is called the *run-length property* of a code. Random sequences also have the property that, if they are shifted by any nonzero number of elements, the resulting sequence will have half its elements the same as in the original sequence and half its elements different from the original sequence. This is called the *shift property* of a code. Following Golomb [4], a deterministic sequence that has the balanced, run length, and shift properties as it grows asymptotically large is referred to as a *pseudorandom sequence*. Because these three properties are often the most important in system analysis, DSSS analysis is often done using random spreading sequences instead of deterministic spreading sequences owing to the former's analytical tractability [5, Chap. 2.2].

Among all linear codes, spreading codes generated from maximal-length sequences, or m-sequences, have many desirable properties. Maximal-length sequences are a type of cyclic code (see Section 8.2.4). Thus, they are generated and characterized by a generator polynomial, and their properties can be derived using algebraic coding theory (see [2, Chap. 3.3] and [5, Chap. 2.2]). These sequences have the maximum period $N = 2^n - 1$ that can be generated by a shift register of length n, so the sequence repeats every NT_c seconds. Moreover, since the sequences are cyclic codes, any time shift of an m-sequence is itself an m-sequence. These sequences also have the property that the modulo-2 addition of an m sequence and a time shift of itself results in a different m-sequence corresponding to a different time shift of the original sequence. This property is called the *shift-and-add* property of m-sequences. The m-sequences have roughly the same number of 0s and 1s over a period: $2^{n-1} - 1$ 0-bits and 2^{n-1} 1-bits. Thus, spreading codes generated from m-sequences, called *maximal linear codes*, have a very small DC component. Moreover, maximal linear codes have approximately the same run-length property as random binary sequences; that is, the number of runs of length r in an n-length sequence is $1/2^r$ for $r < n$ and $1/2^{r-1}$ for $r = n$. Finally, the balanced and shift-and-add properties of m-sequences can be used to show that m sequences have the same shift property as random binary sequences. Therefore, since m sequences have the balanced, run-length, and shift properties of random sequences, they belong to the class of pseudorandom (PN) sequences [5, Chap. 2.2].

The autocorrelation $\rho_c(\tau)$ of a maximal linear spreading code taken over a full period $T = NT_c$ is given by

$$\rho_c(\tau) = \begin{cases} 1 - |\tau|(1 + 1/N)/T_c & |\tau| \le T_c, \\ -1/N & |\tau| > T_c, \end{cases} \tag{13.19}$$

for $|\tau| < (N-1)T_c$, which is illustrated in Figure 13.7. Moreover, since the spreading code is periodic with period $T = NT_c$, the autocorrelation is also periodic with the same period, as shown in Figure 13.8. Thus, if τ is not within a chip time of kNT_c for any integer k, then $\rho_c(\tau) = -1/N = -1/(2^n - 1)$. By making n sufficiently large, the impact of multipath components associated with delays that are not within a chip time of kNT_c can be mostly removed. For delays τ within a chip time of kNT_c, the attenuation is determined by the autocorrelation $\rho_c(\tau)$, which increases linearly as τ approaches kNT_c. The power spectrum of $s_c(t)$ is obtained by taking the Fourier transform of its autocorrelation $\rho_c(\tau)$, yielding

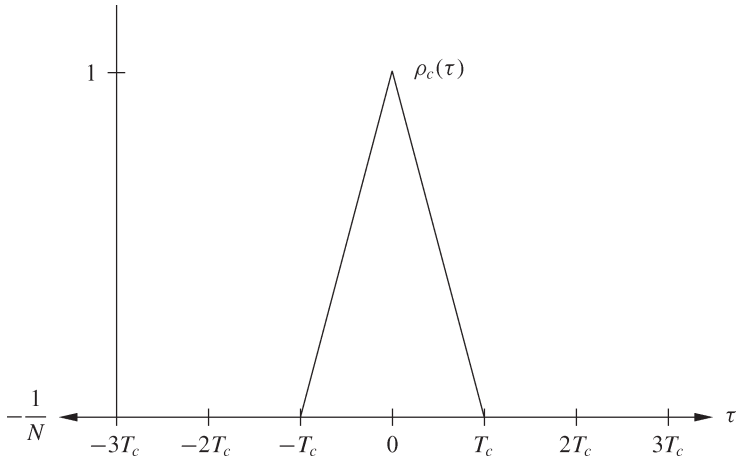


Figure 13.7: Autocorrelation of maximal linear code ($N = T/T_c$).

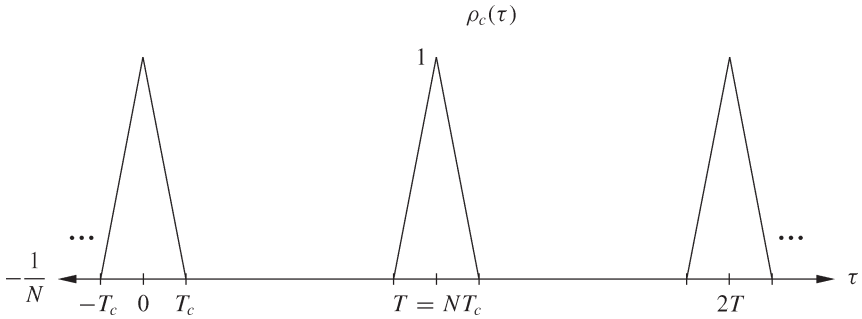


Figure 13.8: Autocorrelation has period $T = NT_c$.

$$P_{s_c}(f) = \sum_{m=-\infty}^{\infty} \frac{N+1}{N^2} \operatorname{sinc}^2\left(\frac{m}{N}\right) \delta\left(f - \frac{m}{T}\right). \tag{13.20}$$

Since $\rho_c(\tau)$ is periodic it follows that $P_{s_c}(f)$ is discrete, with samples every $1/T = 1/NT_c$ Hz.

The periodic nature of the autocorrelation $\rho_c(\tau)$ complicates ISI rejection. In particular, from (13.16), the demodulator associated with the data signal in a spread-spectrum system attenuates the ISI by the autocorrelation $\rho_c(\tau)$ taken over a symbol time T_s. Thus, if the code is designed with $N = T_s/T_c$ chips per symbol, then the demodulator computes the autocorrelation over the full period $T_s = NT_c$ and $\rho_c(\tau)$ is as given in (13.19). Setting $N = T_s/T_c$ is sometimes referred to as a *short spreading code* because the autocorrelation repeats every symbol time, as shown in Figure 13.8 for $T = T_s$. However, short codes exhibit significant ISI from multipath components delayed by approximately an integer multiple of a symbol time, in particular the first few symbols after the desired symbol. If the period of the code is extended so that $N \gg T_s/T_c$, then only multipath at very large delays are not fully attenuated, and these multipath components typically have a low power anyway owing to path loss. Setting $N \gg T_s/T_c$ is sometimes referred to as a *long spreading code*. With long spreading codes the autocorrelation (13.17) performed by the demodulator is taken over a partial period $T = T_s \ll NT_c$ instead of the full period NT_c. The autocorrelation of a maximal linear code over a partial period is no longer characterized by (13.19), so multipath delayed by more than a chip time is no longer attenuated by $-1/N$. Moreover, the partial-period autocorrelation is quite difficult to characterize analytically, since it depends on the starting point in the code over which the partial autocorrelation is taken. By averaging over all starting points, it can be shown that the ISI

attenuation associated with the partial autocorrelation is roughly equal to $1/G$ for G the processing gain, where $G \approx T_s/T_c$, the number of chips per symbol [6, Chap. 9.2].

Although maximal linear codes have excellent properties for ISI rejection, they have a number of properties that make them suboptimal for exploiting the multiuser capabilities of spread spectrum. In particular, there are fewer maximal linear codes of a given length N than in other codes, which limits the number of users who can share the total system bandwidth for multiuser DSSS based on maximal linear codes. Moreover, maximal linear codes generally have relatively poor cross-correlation properties, at least for some sets of codes. In particular, the normalized code cross-correlation can be as high as .37 [6, Chap. 9.2]. Hence for spread-spectrum systems with multiple users, Gold, Kasami, or Walsh codes are used instead of maximal linear codes because they have superior cross-correlation properties. However, these codes can be less effective at ISI rejection than maximal linear codes. More details on these other types of spreading codes will be given in Section 13.4.1.

Example 13.3: Consider a spread-spectrum system using maximal linear codes with period $T = T_s$ and $N = 100$ chips per symbol. Assume the synchronizer has a delay offset of $.5T_c$ relative to the LOS signal component to which it is synchronized. By how much is the power of this signal component reduced by this timing offset?
Solution: For $\tau = .5T_c$ and $N = 100$, the autocorrelation $\rho_c(\tau)$ given by (13.19) is

$$1 - \frac{|\tau|(1 + 1/N)}{T_c} = 1 - \frac{.5T_c(1 + 1/100)}{T_c} = 1 - .5(1.01) = .495. \tag{13.21}$$

Since the signal component is multiplied by $\rho_c(\tau)$, its power is reduced by $\rho_c^2(\tau) = .495^2 = .245 = -6.11$ dB. This is a significant reduction in power, indicating the importance of accurate synchronization, which is discussed in the next section.

13.2.3 Synchronization

We now examine the operation of the synchronizer in Figure 13.5. This operation is separate from carrier phase recovery, so we will assume that the carrier in the demodulator is coherent in phase with the received carrier. The synchronizer must align the timing of the spreading code generator in the receiver with the spreading code associated with one of the multipath components arriving over the channel. A common method of synchronization uses a feedback control loop, as shown in Figure 13.9. The basic premise of the feedback loop is to adjust the delay τ of the spreading code generator until the function $w(\tau)$ reaches its peak value. At this point, under ideal conditions, the spreading code is synchronized to the input, as we now illustrate.

Consider a channel with impulse response $h(t) = \delta(t - \tau_0)$ that introduces a delay τ_0. Neglecting noise, the signal input to the synchronizer from Figure 13.5 is $z(t) = x(t - \tau_0)s_c(t - \tau_0)\cos(2\pi f_c(t - \tau_0))\cos(2\pi f_c t + \phi)$, where ϕ is the phase associated with the carrier recovery loop in the receiver. Assuming perfect carrier recovery, $\cos(2\pi f_c(t - \tau_0)) = \cos(2\pi f_c t + \phi)$, so $z(t) = x(t - \tau_0)s_c(t - \tau_0)\cos^2(2\pi f_c t + \phi)$. The feedback loop will achieve synchronization when $\tau = \tau_0$. We will first assume that $x(t)$ is a BPSK signal with rectangular pulse shaping (so $s_l = \pm\sqrt{E_b}$ and $x(t) = \pm\sqrt{2E_b/T_b}$ is constant over a bit time T_b) and that the spreading codes are maximal linear codes. We will then discuss extensions to more general spreading codes and modulated signals. Assume the spreading codes have period $T = NT_c = T_b$, so their autocorrelation over one period is given by (13.19) and shown in Figure 13.7. Then, setting $\lambda = \sqrt{2/T_b}$, we have

$$w(\tau) = \frac{\lambda}{T}\int_0^T s_l s_c(t-\tau_0)s_c(t-\tau)\cos^2(2\pi f_c t+\phi)\,dt$$
$$\approx \frac{.5\lambda s_l}{T}\int_0^T s_c(t-\tau_0)s_c(t-\tau)\,dt = .5\lambda s_l \rho_c(\tau-\tau_0), \tag{13.22}$$

by (13.18). Since $\rho_c(\tau-\tau_0)$ reaches its maximum at $\tau-\tau_0=0$ and since $s_l=\pm\sqrt{E_b}$, the feedback control loop will adjust τ such that $|w(\tau)|$ increases. In particular, suppose $|\tau-\tau_0|>T_c$. Then, by (13.19), $\rho_c(\tau-\tau_0)=-1/N$ and the synchronizer is operating outside the triangular region of the autocorrelation function shown in Figure 13.7. The feedback control loop will therefore adjust τ, typically in increments of T_c, until $|w(\tau)|$ increases above $-1/N$. This increase occurs when τ is sufficiently adjusted that $|\tau-\tau_0|<T_c$. At this point the synchronizer is within a chip time of perfect synchronization, which is sometimes referred to as *coarse synchronization* or *acquisition*. In general the channel has many multipath components, in which case coarse synchronization will synchronize to the first multipath component it finds above a given power threshold.

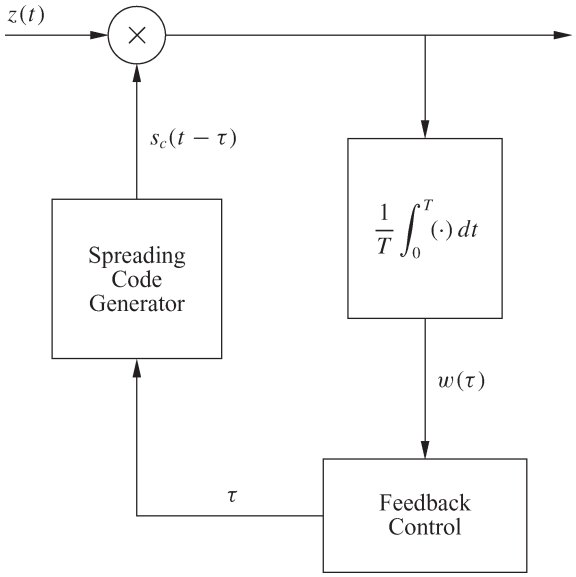


Figure 13.9: Synchronization loop for DSSS.

An alternative to the feedback control loop for acquisition is a parallel-search acquisition system. This system has multiple branches that correlate the received signal against a delayed version of the spreading code, where each branch has a different delay equal to an integer multiple of the chip time. The synchronization locks to the branch with the maximum correlator output. A similar structure is used in a RAKE receiver, discussed in the next section, to coherently combine multipath components at different delays. For both synchronization methods, the coarse acquisition often uses short codes with a small period T to reduce acquisition time. If long codes are used, the acquisition time can be shortened by performing the integration in the feedback loop over a fraction of the entire code period. In this case, as long as the partial autocorrelation is small for delays greater than a chip time and is above a given threshold for delays within a chip time, the acquisition loop can compare the partial autocorrelation against the threshold to determine if coarse acquisition has occurred. For the fine tuning that follows coarse acquisition, long codes with integration over the full period are typically used to make the synchronization as precise as possible.

Once coarse synchronization is achieved, the feedback control loop makes small adjustments to τ to fine-tune its delay estimate such that $\tau \approx \tau_0$. This is called *fine synchronization* or *tracking*. Suppose through course synchronization we obtain $\tau - \tau_0 = T_c$. Referring to Figure 13.7, this implies that the synchronizer is operating on the far right edge of the triangular correlation function. As τ is further decreased, $\tau - \tau_0$ decreases toward zero, and the synchronization "walks backwards" toward the peak of the autocorrelation at $\tau - \tau_0 = 0$. Once the peak is attained, the synchronizer locks to the delay τ_0. Because of the time-varying nature of the channel as well as interference, multipath, and noise, τ must be adjusted continuously to optimize synchronization under these dynamic operating conditions. Spread-spectrum tracking often uses the same timing recovery techniques discussed in Section 5.6.3 for narrowband systems.

The acquisition and tracking procedures for more general spreading codes are very similar. Since all periodic spreading codes have an autocorrelation that peaks at zero, the course and fine synchronization will adjust their estimate of the delay to try to maximize the autocorrelation output of the integrator. Synchronization performance is highly dependent on the shape of the autocorrelation function. A sharp autocorrelation facilitates accurate fine tuning of the synchronization. Noise, fading, interference, and ISI will complicate both coarse and fine synchronization, since the output of the integrator in Figure 13.9 will be distorted by these factors.

When $s(t)$ is not binary and/or it has a symbol time that is less than the code period, the integrator output will depend on the data symbol(s) over the duration of the integration. This is the same situation as in carrier and timing recovery of narrowband systems with unknown data, discussed in Section 5.6, and similar techniques can be applied in this setting. Note that we have also neglected carrier phase recovery in our analysis, assuming that the receiver has a carrier recovery loop to obtain a coherent phase reference on the received signal. Carrier recovery techniques were discussed in Section 5.6, but these techniques must be modified for spread-spectrum systems because the spreading codes affect the carrier recovery process [7]. Synchronization is a challenging aspect of spread-spectrum system design, especially in time-varying wireless environments. Much work has been devoted to developing and analyzing spread-spectrum synchronization techniques. Details on the main techniques and their performance can be found in [2, Chaps. 4–5], [8, Chap. 6], [9, Part 4:1–2], and [10, Chap. 12.5].

13.2.4 RAKE Receivers

The spread-spectrum receiver shown in Figure 13.5 will synchronize to one of the multipath components in the received signal. The multipath component to which it is synchronized is typically the first one acquired during the coarse synchronization that is above a given threshold. This may not be the strongest multipath component; moreover, the result is that all other multipath components are then treated as interference. A more complicated receiver can have several branches, with each branch synchronized to a different multipath component. This structure is called a *RAKE* receiver,[4] and it typically assumes there is a multipath component at each integer multiple of a chip time. Thus, the time delay of the spreading code between branches is T_c, as shown in Figure 13.10. Note that the carrier phase offset associated with each multipath component depends on its delay: this offset is determined by the coherent demodulator in each branch. Alternatively, each branch can be downconverted to baseband prior to multiplication with the spreading code, in which case each branch performs carrier phase recovery prior to downconversion. The RAKE is essentially another form of diversity combining, since the spreading code induces a path diversity on the transmitted signal so that independent multipath components separated by more than a chip time can be resolved. Any of the combining techniques discussed in Chapter 7 may be used, although in practice MRC is the most common.

[4]The name RAKE comes from the notion that the multibranch receiver resembles a garden rake and has the effect of raking up the energy associated with the multipath components on each of its branches. The RAKE was invented in the 1950s to deal with the ionospheric multipath on a spread-spectrum high-frequency transcontinental link. The name was coined by the RAKE inventors, Paul Green and Bob Price.

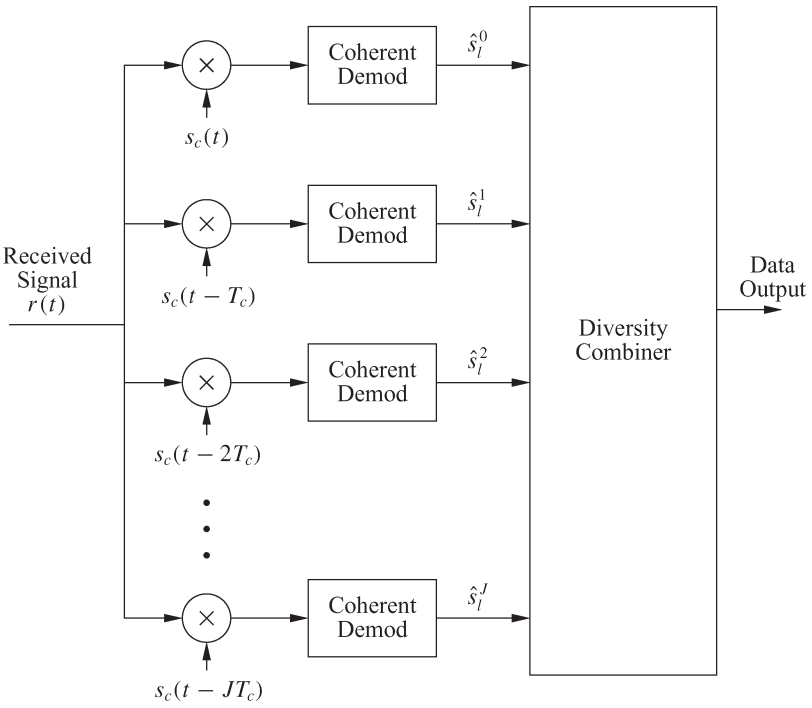


Figure 13.10: RAKE receiver.

In order to study the behavior of RAKE receivers, assume a channel model with impulse response $h(t) = \sum_{j=0}^{J} \alpha_j \delta(t - jT_c)$, where α_j is the gain associated with the jth multipath component. This model, described in Section 3.4, can approximate a wide range of multipath environments by matching the statistics of the complex gains to those of the desired environment. The statistics of the α_j have been characterized empirically in [11] for outdoor wireless channels. With this model, each branch of the RAKE receiver in Figure 13.10 synchronizes to a different multipath component and coherently demodulates its associated signal. A larger J implies a higher receiver complexity but also increased diversity. Then, from (13.14) and (13.15), the output of the ith branch demodulator is

$$\hat{s}_l^i = \alpha_i s_l + \sum_{\substack{j=1 \\ i \neq j}}^{J-1} \alpha_j \rho_c(iT_c - jT_c) s_{lj} + n_l^i, \tag{13.23}$$

where s_l is the symbol transmitted over symbol time $[lT_s,\ (l+1)T_s]$ (i.e., the symbol associated with the LOS path); s_{lj} is the symbol transmitted over symbol time $[lT_s - jT_c,\ (l+1)T_s - jT_c]$, which we assume to be constant; and n_l^i is the noise sample associated with the demodulator in the ith branch. If s_{lj} is not constant over $[lT_s - jT_c,\ (l+1)T_s - jT_c]$ then the ISI term in (13.23) is more complicated and involves partial autocorrelations. However, in all cases the ISI is reduced by roughly the autocorrelation $\rho_c((i-j)T_c)$. The diversity combiner coherently combines the demodulator outputs. In particular, with selection combining the branch output $\hat{s}_l^i$ with the largest path gain a_i is output from the combiner, with equal gain combining all demodulator outputs are combined with equal weighting, and with maximal ratio combining the demodulator outputs are combined with a weight equal to the branch SNR – or to the SINR (signal-to-interference-plus-noise power ratio) if ISI interference is taken into account. If $\rho_c(\tau) \approx 0$ for $|\tau| > T_c$ then we can neglect the ISI terms in each branch, and the performance of the RAKE receiver with J branches is identical to any other J-branch diversity technique. A comprehensive study

of RAKE performance for empirically derived channel models was made by Turin in [11].

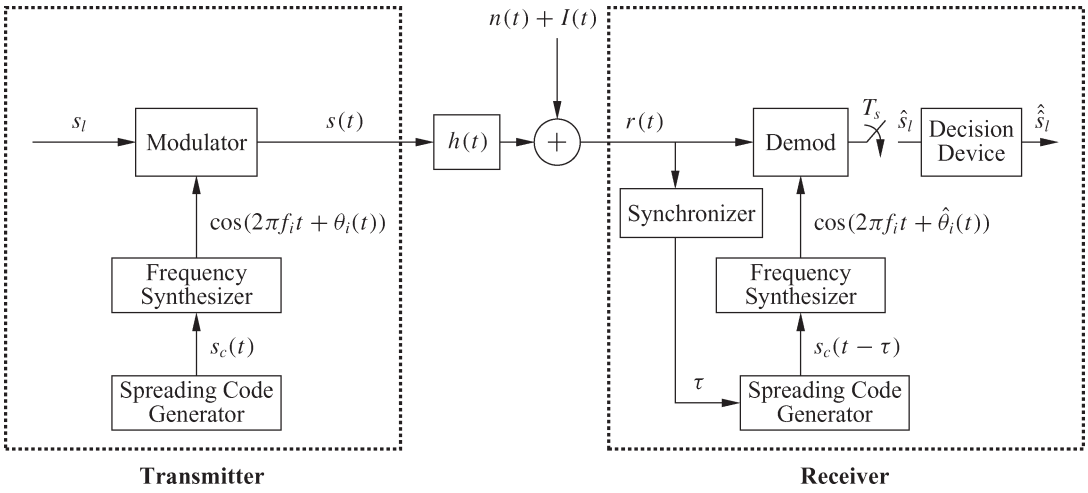


Figure 13.11: FHSS system model.

Spread spectrum is not usually used for diversity alone, since it requires significantly more bandwidth than other diversity techniques. However, if spread-spectrum signaling is chosen for its other benefits – such as its multiuser or interference rejection capabilities – then RAKEs provide a simple mechanism to obtain diversity benefits.

13.3 Frequency-Hopping Spread Spectrum (FHSS)

An end-to-end frequency-hopping spread-spectrum system is illustrated in Figure 13.11. The spreading code is input to the frequency synthesizer in order to generate the hopping carrier signal $c(t) = \cos(2\pi f_i t + \theta_i(t))$, which is then input to the modulator to upconvert the modulated signal to the carrier frequency. The modulator can be coherent, noncoherent, or differentially coherent, although coherent modulation is not as common as noncoherent modulation owing to the difficulties in maintaining a coherent phase reference while hopping the carrier over a wide bandwidth [9, Part 2:2]. At the receiver, a synchronizer is used to synchronize the locally generated spreading code to that of the incoming signal. Once synchronization is achieved, the spreading code is input to the frequency synthesizer to generate the hopping pattern of the carrier, which is then input to the demodulator for downconversion. For a noncoherent or differentially coherent modulator, it is not necessary to synchronize the phase associated with the receive carrier to that of the transmit carrier.

As with DSSS, the synchronization procedure for FH systems is typically performed in two stages. First, a coarse synchronization is done to align the receiver hop sequence to within a fraction of the hop duration T_c associated with the transmitted FH signal. The process is similar to the coarse synchronization of DSSS: the received FH signal plus noise is correlated with the local hopping sequence by multiplying the signals together and computing the energy in their product. If this energy exceeds a given threshold, coarse acquisition is obtained; otherwise, the received FH signal is shifted in time by T_c and the process is repeated. Coarse acquisition can also be done in parallel using multiple hop sequences, each shifted in time by a different integer multiple of T_c. Once coarse acquisition is obtained, fine tuning occurs by continually adjusting the timing of the frequency hopper to maximize the correlation between the receiver hopping sequence and the received signal. More details on FH synchronization and an analysis of system performance under synchronization errors can be found in [9, Part 4].

The impact of multipath on FH systems was discussed in Section 13.1. There we saw that an FH system

does not exhibit fading if the multipath components have delay exceeding the hop time, since only one nonfading signal component arrives during each hop. When multipath does cause flat or frequency-selective fading, the performance analysis is the same as for a slowly time-varying non-hopping system. However, the impact of narrowband interference on FH systems, as characterized by the probability of symbol error, is more difficult to determine. In fact, this error probability depends on the exact structure of the interfering signal and how it affects the specific modulation in use, as we now describe.

We will focus on symbol error probability for an SFH system without coding, where the interference, if present, is constant over a symbol time. The analysis for fast frequency hopping is more complicated because interference changes over a symbol time, making it more difficult to characterize its statistics and the resulting impact on the symbol error probability. So consider an SFH system with M out of the N frequency bands occupied by a narrowband interferer. Assuming the signal hops uniformly over the entire frequency band, the probability of any given hop being in the same band as an interferer is then M/N. The probability of symbol error is obtained by conditioning on the presence of an interferer over the given symbol period:

$$\begin{aligned} P_s &= p(\text{symbol error} \mid \text{no interference})p(\text{no interference}) \\ &\quad + p(\text{symbol error} \mid \text{interference})p(\text{interference}) \\ &= \frac{N-M}{N}p(\text{symbol error} \mid \text{no interference}) \\ &\quad + \frac{M}{N}p(\text{symbol error} \mid \text{interference}). \end{aligned} \tag{13.24}$$

In the absence of interference, the probability of symbol error just equals that of the modulated data signal transmitted over an AWGN channel with received SNR γ_s, which we will denote as P_s^{AWGN}. Note that γ_s is the received SNR at the input to the demodulator in the absence of interference, so multipath components removed in the despreading process do not affect this SNR. However, γ_s will be affected by the channel gain at the carrier frequency for the multipath components that are not removed by despreading. For most coherent modulations, $P_s^{\text{AWGN}} \approx \alpha_M Q(\sqrt{\beta_M \gamma_s})$ for α_M and β_M dependent on the modulation, as discussed in Section 6.1.6. The P_s^{AWGN} for noncoherent or differentially coherent modulations in AWGN are generally more complex [12, Chap. 1.1]. Given P_s^{AWGN}, it remains only to characterize the probability of error when interference is present, $p(\text{symbol error} \mid \text{interference})$, in order to determine P_s in (13.24). If we denote this probability as P_s^{INT}, then (13.24) becomes

$$P_s = \frac{N-M}{N}P_s^{\text{AWGN}} + \frac{M}{N}P_s^{\text{INT}}. \tag{13.25}$$

Let us now examine P_s^{INT} more closely. This symbol error probability will depend on the exact characteristics of the interference signal. Consider first a narrowband interferer with the same statistics as AWGN within the bandwidth of the modulated signal; an interferer with these characteristics is sometimes referred to as a *partial band noise jammer*. For this type of interferer, P_s^{INT} is obtained by treating the interference as an additional AWGN component with power N_J within the bandwidth of the modulated signal. The total noise power is then $N_0B + N_J$ and so the SINR in the presence of this interference becomes

$$\gamma_s^{\text{INT}} = \gamma_s \frac{N_0 B}{N_0 B + N_J}, \tag{13.26}$$

which yields

$$P_s^{\text{INT}} = P_s^{\text{AWGN}}(\gamma_s^{\text{INT}}). \tag{13.27}$$

Suppose now that the interference consists of a tone at the hopped carrier frequency with some offset phase. Then the demodulator output $\hat{s}_l$ in Figure 13.11 is given by

$$\hat{s}_l = a_l s_l + n_l + I_l, \tag{13.28}$$

where a_l is the channel gain associated with the received signal after despreading, n_l is the AWGN sample, and $I_l = \sqrt{I}e^{j\phi_l}$ is the interference term with phase offset ϕ_l. Note that, since this is a wideband channel, fading is frequency selective and so (i) the channel gain a_l will depend on the carrier frequency and (ii) some hops may be associated with very poor channel gains. The impact of the additional interference term I_l will depend on the modulation. For example, with coherent MPSK and assuming $\angle s_l = 0$,

$$P_s = 1 - p(|\angle(a_l s_l + n_l + I_l)| \leq \pi/M). \tag{13.29}$$

In general, computing P_s for either coherent or noncoherent modulation requires finding the distribution of the random phase $\angle(n_l + I_l)$. This distribution and the resulting P_s are derived in [9, Parts 2–3] for noncoherent, coherent, and differentially coherent modulations for a number of different interference models. Coding or coding with interleaving is often used in FH systems to compensate for frequency-selective fading as well as narrowband interference jamming. Analysis of coded systems with interference can be found in [9, Part 2:2].

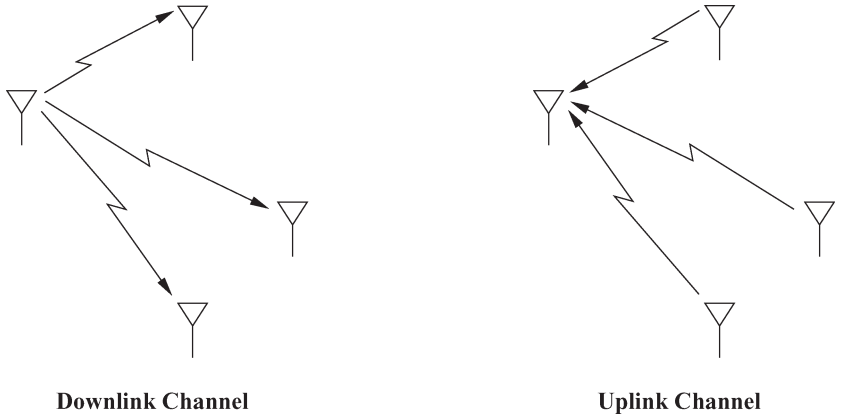


Figure 13.12: Downlink and uplink channels.

13.4 Multiuser DSSS Systems

Spread spectrum can also be used as a mechanism whereby many users can share the same spectrum. Using spreading code properties to support multiple users within the same spread bandwidth is also called spread-spectrum multiple access (SSMA), which is a special case of code-division multiple access (CDMA). In multiuser spread spectrum, each user is assigned a unique spreading code or hopping pattern, which is used to modulate its data signal. The transmitted signals for all users are superimposed in time and in frequency. The spreading codes or hopping patterns can be orthogonal, in which case users do not interfere with each other under ideal propagation conditions, or they can be nonorthogonal, in which case interference between users does exist but is reduced by the spreading code properties. Thus, whereas spread spectrum for single-user systems is spectrally inefficient because it uses more bandwidth than the minimum needed to convey the information signal, spread-spectrum multiuser systems can support an equal or larger number of users in a given bandwidth than other forms of spectral sharing such as time division or frequency division. However, if the spreading mechanisms are nonorthogonal – either by design

or through channel distortion – then users interfere with each other. If there is too much interference between users, the performance of all users degrades. Comparison of the spectral efficiency for different spectral-sharing methods in multiuser and cellular systems will be discussed in Chapters 14 and 15.

Performance of multiuser spread spectrum also depends on whether the multiuser system is a *downlink* channel (one transmitter to many receivers) or an *uplink* channel (many transmitters to one receiver). These channel models are illustrated in Figure 13.12: the downlink channel is also called a broadcast channel or forward link, and the uplink channel is also called a multiple access channel or reverse link. The performance differences of DSSS in uplink and downlink channels result from the fact that, in the downlink, all transmitted signals are typically synchronous because they originate from the same transmitter. Moreover, both the desired signal and interference signals pass through the same channel before reaching the desired receiver. In contrast, users in the uplink channel are typically asynchronous, since they originate from transmitters at different locations, and the transmitted signals of the users travel through different channels before reaching the receiver. In this section we will analyze the multiuser properties of DSSS for both downlinks and uplinks. In Section 13.5 we treat multiuser FHSS systems.

13.4.1 Spreading Codes for Multiuser DSSS

Multiuser DSSS is accomplished by assigning each user a unique spreading code sequence $s_{c_i}(t)$. As described in Section 13.2.2, the autocorrelation function of the spreading code determines its multipath rejection properties. The cross-correlation of different spreading codes determines the amount of interference between users whose signals are modulated by these codes. The signals of asynchronous users arrive at the receiver with arbitrary relative delay τ, and the cross-correlation of the codes assigned to user i and user j over one symbol time with this delay is given by

$$\rho_{ij}(t,\ t-\tau) = \frac{1}{T_s}\int_0^{T_s} s_{c_i}(t)s_{c_j}(t-\tau)dt = \frac{1}{N}\sum_{n=1}^{N} s_{c_i}(nT_c)s_{c_j}(nT_c-\tau), \tag{13.30}$$

where N is the number of chips per symbol time. As with the autocorrelation, if $s_{c_i}(t)$ and $s_{c_j}(t)$ are periodic with period T_s then the cross-correlation depends only on the time difference τ, so $\rho_{ij}(t,\ t-\tau) \triangleq \rho_{ij}(\tau)$. For synchronous users, their signals arrive at the receiver aligned in time, so $\tau = 0$ and the cross-correlation becomes

$$\rho_{ij}(0) = \frac{1}{T_s}\int_0^{T_s} s_{c_i}(t)s_{c_j}(t)\,dt = \frac{1}{N}\sum_{n=1}^{N} s_{c_i}(nT_c)s_{c_j}(nT_c). \tag{13.31}$$

Ideally, since interference between users is dictated by the cross-correlation of the spreading code, for $i \neq j$ we would like $\rho_{ij}(\tau) = 0$ for all τ for asynchronous users and $\rho_{ij}(0) = 0$ for synchronous users; this would eliminate interference between users. A set of spreading codes with this *cross-correlation property* is called an *orthogonal* code set, and a set of spreading codes that does not satisfy this property is called a *nonorthogonal* code set. It is not possible to obtain orthogonal codes for asynchronous users, and for synchronous users there is only a finite number of spreading codes that are orthogonal within any given bandwidth. Thus, an orthogonality requirement restricts the number of different spreading codes (and the corresponding number of users) in a synchronous DSSS multiuser system. We now describe the most common chip sequences and their associated spreading codes that are used in multiuser DSSS systems.

GOLD CODES

Gold codes have worse autocorrelation properties than maximal-length codes but better cross-correlation properties if properly designed. The chip sequences associated with a Gold code are produced by the binary addition of two m-sequences each length $N = 2^n - 1$, and they inherit the balanced, run-length, and shift properties of these

component sequences (hence they are pseudorandom sequences). Gold codes take advantage of the fact that, if two distinct m-sequences with time shifts τ_1 and τ_2 are added together modulo 2, then the resulting sequence is unique for every unique value of τ_1 or τ_2. Thus, a very large number of unique Gold codes can be generated, which allows for a large number of users in a multiuser system. However, if the m-sequences that are modulo-2 added to produce a Gold code are chosen at random, then the cross-correlation of the resulting code may be quite poor. Thus, Gold codes are generated by the chip sequences associated with the modulo-2 addition *of preferred* pairs of m-sequences. These preferred pairs are chosen to obtain good cross-correlation properties in the resulting Gold code. However, the preferred pairs of m-sequences have different autocorrelation properties than general m-sequences. A method for choosing the preferred pairs such that the cross-correlation and autocorrelation functions of the resulting Gold code are bounded was given by Gold in [13] and can also be found in [6, Chap. 9.2], [8, Apx. 7], and [14]. The preferred sequences are chosen so that Gold codes have a three-valued cross-correlation over a code period with values

$$\rho_{ij}(\tau) = \begin{cases} -1/N, \\ -t(n)/N, \\ (t(n)-2)/N, \end{cases} \tag{13.32}$$

where

$$t(n) = \begin{cases} 2^{(n+1)/2}+1 & n \text{ odd}, \\ 2^{(n+2)/2}+1 & n \text{ even}. \end{cases} \tag{13.33}$$

The autocorrelation takes on the same three values.

KASAMI CODES

Kasami sequences have properties that are similar to the preferred sequences used to generate Gold codes, and they are also derived from m-sequences. However, the Kasami codes have better cross-correlation properties than Gold codes. There are two different sets of Kasami sequences that are used to generate Kasami codes, the large set and the small set. To generate the small set, we begin with an m-sequence a of length $N = 2^n - 1$ for n even and form a new shorter sequence a' by sampling every $2^{n/2}+1$ elements of a. The resulting sequence a' will have period $2^{n/2}-1$. We then generate a small set of Kasami sequences by taking the modulo-2 sum of a with all cyclic shifts of the a' sequence. There are $2^{n/2}-2$ such cyclic shifts and – by also including the original sequence a as well as $\mathrm{mod}_2(a+a')$ in the set – we obtain a set of $2^{n/2}$ binary sequences of length 2^n-1. As with the Gold codes, the auto-correlation and cross-correlation of the Kasami spreading codes obtained from the Kasami sequences over a code period are three-valued, taking on the values

$$\rho_{ij}(\tau) = \begin{cases} -1/N, \\ -s(n)/N, \\ (s(n)-2)/N, \end{cases} \tag{13.34}$$

where $s(n) = 2^{n/2}+1$. Since $|s(n)| < |t(n)|$, Kasami codes have better autocorrelation and cross-correlation than Gold codes. In fact, Kasami codes achieve the Welch lower bound for autocorrelation and cross-correlation for any set of $2^{n/2}$ sequences of length 2^n-1 and hence are optimal (in terms of minimizing the autocorrelation and cross-correlation) for any such code (see [9, Part 1:5] and [14]) .

The large set of Kasami sequences is formed in a similar fashion. It has a larger number of sequences than the smaller set and thus can support more users in a multiuser system, but the autocorrelation and cross-correlation properties across the spreading codes generated from this larger set are inferior to those generated from the smaller set. To obtain the large set, we take an m-sequence a of length $N = 2^n - 1$ for n even and form two new sequences

a' and a'' by sampling the original sequence every $2^{n/2}+1$ elements for a' and every $2^{(n+2)/2}+1$ elements for a''. The set is then constituted by adding a, a', and a'' for all cyclic shifts of a' and a''. The number of such sequences is $2^{3n/2}$ if n is a multiple of 4 and is $2^{3n/2}+2^{n/2}$ if $\mathrm{mod}_4(n)=2$. The autocorrelation and cross-correlation of the spreading codes generated from this set over a code period can take on one of five values:

$$\rho(\tau)=\begin{cases}-1/N,\\(-1\pm 2^{n/2})/N,\\(-1\pm(2^{n/2}+1))/N.\end{cases} \tag{13.35}$$

Since these values exceed those for codes generated from the small Kasami set, we see that codes generated from the large Kasami set have inferior cross-correlation and autocorrelation properties to those generated from the small Kasami set.

Example 13.4: Find the number of sequences and the magnitude of the worst-case cross-correlation over a code period for spreading codes generated from the small and large Kasami sequences with $n=10$.
Solution: For the small set, there are $2^{n/2}=2^5=32$ sequences. From (13.34), the largest-magnitude cross-correlation of the corresponding Kasami code is

$$\frac{1}{N}[2^{n/2}+1]=\frac{1}{2^{10}-1}[2^5+1]=.032. \tag{13.36}$$

For the large set, $\mathrm{mod}_4(10)=2$, so there are $2^{3n/2}+2^{n/2}=2^{15}+2^{10}=33792$ sequences, which is three orders of magnitude more codes than in the small set. The largest-magnitude spreading code cross-correlation is thus

$$\frac{1}{N}[2^{n/2}+2]=\frac{1}{2^{10}-1}[2^5+2]=.033. \tag{13.37}$$

So there is a slightly larger cross-correlation, the price paid for the significant increase in the number of codes.

WALSH–HADAMARD CODES

Walsh–Hadamard codes of length $N=T_s/T_c$ that are synchronized in time are orthogonal over a symbol time, so that the cross-correlation of any two codes is zero. Thus, synchronous users modulated with Walsh–Hadamard codes can be separated out at the receiver with no interference between them, as long as the channel does not corrupt the orthogonality of the codes. (Delayed multipath components are not synchronous with the LOS paths, and thus the multipath components associated with different users will cause interference between users. The loss of orthogonality can be quantified by the orthogonality factor [15]). Although it is possible to synchronize users on the downlink, where all signals originate from the same transmitter, it is more challenging to synchronize users in the uplink because they are not co-located. Hence, Walsh–Hadamard codes are rarely used for DSSS uplink channels. Walsh–Hadamard sequences of length N are obtained from the rows of an $N\times N$ *Hadamard matrix* $\mathbf{H}_N$. For $N=2$ the Hadamard matrix is

$$\mathbf{H}_2=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}. \tag{13.38}$$

Larger Hadamard matrices are obtained using $\mathbf{H}_2$ and the recursion

$$\mathbf{H}_{2N}=\begin{bmatrix}\mathbf{H}_N & \mathbf{H}_N\\ \mathbf{H}_N & -\mathbf{H}_N\end{bmatrix}. \tag{13.39}$$

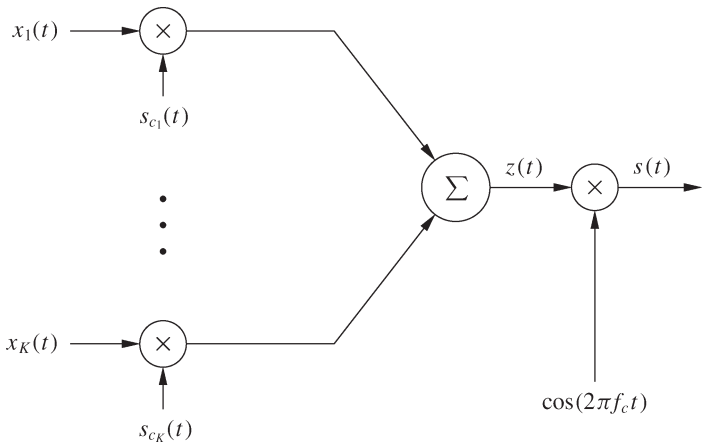


Figure 13.13: Downlink transmitter.

Each row of $\mathbf{H}_N$ specifies a different chip sequence, so the number of spreading codes in a Walsh–Hadamard code is N. Thus, DSSS with Walsh–Hadamard codes can support at most $N = T_s/T_c$ users. Since DSSS uses roughly N times more bandwidth than required for the information signal, approximately the same number of users could be supported by dividing up the total system bandwidth into N nonoverlapping channels (frequency division). Similarly, the same number of users can be supported by dividing time up into N orthogonal timeslots (time division), where each user operates over the entire system bandwidth during his timeslot. Hence, all multiuser techniques that assign orthogonal channels to preclude interference will accommodate approximately the same number of users.

The performance of a DSSS multiuser system depends both on the spreading code properties as well as the channel over which the system operates. In the next section we will study performance of DSSS multiuser systems over downlinks, and performance over uplinks will be treated in Section 13.4.3.

13.4.2 Downlink Channels

The transmitter for a DSSS downlink system is shown in Figure 13.13, with the channel and receiver shown in Figure 13.14. We focus on analyzing the in-phase signal component; a similar analysis applies for the quadrature component. In the downlink the signals of all users are typically sent simultaneously by the transmitter (base station), and each receiver must demodulate its individual signal. Thus we can assume that all signals are synchronous, which allows the use of orthogonal spreading codes such as the Walsh–Hadamard codes. However, the use of orthogonal codes limits the number of users the downlink can support, so such codes are not always used.

Consider a K-user system, where the transmitter sends to K independent users. The base-band modulated signal associated with the kth user is

$$x_k(t) = \sum_l s_{kl} g(t - lT_s), \tag{13.40}$$

where $g(t) = \sqrt{2/T_s}$ is the pulse shape (assumed rectangular), T_s is the symbol time, and s_{kl} is the kth user's symbol over the lth symbol time. The transmitter consists of K branches, where the kth branch multiplies the kth user's signal $x_k(t)$ with the spreading code $s_{c_k}(t)$. The branches are summed together, resulting in the baseband multiuser signal

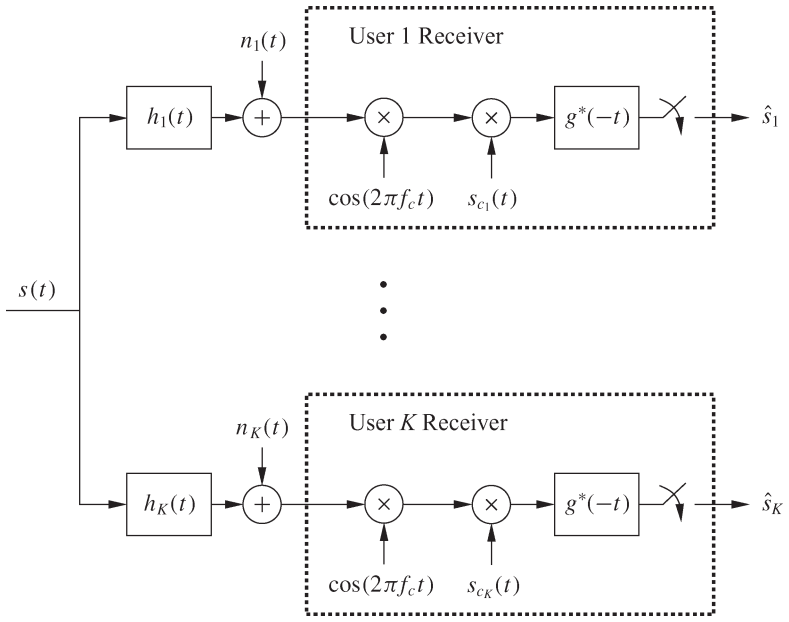


Figure 13.14: Downlink channel and receiver.

$$z(t) = \sum_{k=1}^{K} x_k(t) s_{c_k}(t) = \sum_{k=1}^{K} \sqrt{\frac{2}{T_s}} s_{kl} s_{c_k}(t). \tag{13.41}$$

This multiuser signal is multiplied by the carrier to obtain the passband signal $s(t)$, which is transmitted over the channel.

The signal received by user k first passes through its channel, which has impulse response $h_k(t)$ and AWGN. Thus, the received signal at the kth user's receiver is $s(t) * h_k(t) + n_k(t)$. This signal is downconverted, and we assume perfect carrier recovery so that the phase ϕ_k of the carrier in user k's receiver is perfectly matched to that of its incoming signal. Thus we can normalize the delay of the kth user's LOS (or minimum delay) path and its associated phase shift to zero. After downconversion, the signal is then multiplied by the kth user's spreading code $s_{c_k}(t)$, which is assumed to be perfectly synchronized to the kth user's spreading code in the received signal.[5] The signal is then baseband demodulated via a matched filter (i.e., it is multiplied by $\sqrt{2/T_s}$ and integrated over a symbol time). The demodulator output is sampled every T_s to obtain an estimate of the symbol transmitted by the kth user over that symbol time. Comparing Figure 13.5 and Figure 13.14, we see that the kth user's receiver is identical to the matched filter detector in a single-user DSSS system. Thus, in the absence of multiuser interference, the kth user has identical performance as in a single-user DSSS system. However, when multiuser interference is taken into account, the demodulator output includes components associated with the kth user's signal, interference terms from other users' signals, and noise. In particular, the demodulator output associated with the kth user over the lth symbol time is given by

[5]This synchronization is even more difficult than in the single-user case, since here it must be done in the presence of multiple spread signals. In fact, some spreading code sets are obtained by shifting a single spreading code by some time period. For these systems there must be a control channel to inform the receiver which time shift corresponds to its desired signal. More details on the synchronization for these systems can be found in [8, Chap. 6].

$$
\begin{aligned}
\hat{s}_k &= \sqrt{\frac{2}{T_s}} \int_0^{T_s} [s(t) * h_k(t) + n_k(t)] s_{c_k}(t) \cos(2\pi f_c t)\, dt \\
&= \sqrt{\frac{2}{T_s}} \int_0^{T_s} [z(t) * h_k^{\mathrm{LP}}(t)] s_{c_k}(t) \cos^2(2\pi f_c t) dt + \sqrt{\frac{2}{T_s}} \int_0^{T_s} n_k(t) s_{c_k}(t) \cos(2\pi f_c t)\, dt \\
&= \frac{2}{T_s} \int_0^{T_s} \left[\sum_{j=1}^{K} s_{jl} s_{c_j}(t) * h_k^{\mathrm{LP}}(t) \right] s_{c_k}(t) \cos^2(2\pi f_c t)\, dt \\
&\quad + \sqrt{\frac{2}{T_s}} \int_0^{T_s} n_k(t) s_{c_k}(t) \cos(2\pi f_c t)\, dt \\
&= \frac{2}{T_s} \int_0^{T_s} [s_{kl} s_{c_k}(t) * h_k^{\mathrm{LP}}(t)] s_{c_k}(t) \cos^2(2\pi f_c t)\, dt \\
&\quad + \frac{2}{T_s} \int_0^{T_s} \left[\sum_{\substack{j=1 \\ j \neq k}}^{K} s_{jl} s_{c_j}(t) * h_k^{\mathrm{LP}}(t) \right] s_{c_k}(t) \cos^2(2\pi f_c t)\, dt \\
&\quad + \sqrt{\frac{2}{T_s}} \int_0^{T_s} n_k(t) s_{c_k}(t) \cos(2\pi f_c t)\, dt,
\end{aligned} \tag{13.42}
$$

where $h_k^{\mathrm{LP}}(t)$ is the in-phase component of the equivalent lowpass filter for $h_k(t)$, s_{kl} is the kth user's transmitted symbol over the lth symbol period that is being recovered, and s_{jl} is the transmitted symbol of the jth user over this symbol period, which causes interference. Note that (13.42) consists of three separate terms. The first term corresponds to the received signal of the kth user alone, the second term represents interference from other users in the system, and the last term is the AWGN sample, which we denote as n_k. The first term and the noise sample are characterized by the analysis in Section 13.2 for single-user systems. The second term depends on both the channel $h_k^{\mathrm{LP}}(t)$ and the spreading code properties, as we now show.

To examine the characteristics of the multiuser interference, let us assume that the kth user's channel has a normalized delay of zero; that is, $h_k^{\mathrm{LP}}(t) = \alpha_k \delta(t)$ for some real number α_k. Then (13.42) becomes

$$
\begin{aligned}
\hat{s}_k &= \frac{2}{T_s} \int_0^{T_s} \alpha_k s_{kl} s_{c_k}^2(t) \cos^2(2\pi f_c t)\, dt \\
&\quad + \frac{2}{T_s} \int_0^{T_s} \sum_{\substack{j=1 \\ j \neq k}}^{K} \alpha_k s_{jl} s_{c_j}(t) s_{c_k}(t) \cos^2(2\pi f_c t)\, dt + n_k \\
&\approx \alpha_k s_{kl} + \alpha_k \sum_{\substack{j=1 \\ j \neq k}}^{K} s_{jl} \rho_{jk}(0) + n_k,
\end{aligned} \tag{13.43}
$$

where $\rho_{jk}(0)$ is the cross-correlation between $s_{c_k}(t)$ and $s_{c_j}(t)$ for a timing offset of zero, since the users are assumed to be synchronous.[6] We define

[6] If the users were not synchronous, which is unusual in a downlink channel, then the cross-correlation $\rho_{jk}(0)$ in (13.43) would be replaced by $\rho_{jk}(\tau_{jk})$ for $\tau_{jk}$the relative delay between the received signal from users j and k. This assumes s_{jl} is constant over the integration; if not, the interference term depends on the different symbol values over the integration.

$$I_{kl} = \alpha_k \sum_{\substack{j=1 \\ j \neq k}}^{K} s_{jl} \rho_{jk}(0) \tag{13.44}$$

as the multiuser interference to the kth user at the demodulator output. We see from (13.43) that the kth user's symbol s_{kl} is attenuated by the channel gain but not affected by the spreading and despreading, exactly as in the single-user case. The noise sample n_k is also the same as in a single-user non-spread system. The interference from other users is attenuated by the kth user's channel gain α_k and the cross-correlation of the codes $\rho_{jk}(0)$. For orthogonal (e.g., Walsh–Hadamard) codes, $\rho_{jk}(0) = 0$ so there is no interference between users. For nonorthogonal codes, $\rho_{jk}(0)$ depends on the specific codes assigned to users j and k; for example, with Gold codes, $\rho_{jk}(0)$ can take on one of three possible values. Note that both the kth user's signal and the interference are attenuated by the same channel gain a_k, since both signal and interference follow the same path from the transmitter to the receiver. As we will see in the next section, this is not the case for DSSS uplink systems.

If the interference in a multiuser system has approximately Gaussian statistics then we can treat it as an additional noise term and determine system performance based on the SINR for each user. However, the Gaussian approximation is often inaccurate, even when the number of interferers is large [16]. Moreover, in fading the interference terms are correlated, since they all experience the same fading α_k. Thus, the interference can only be approximated as conditionally Gaussian, conditioned on the fading. The conditionally Gaussian approximation is most accurate when the number of interferers is large, since the sum of a large number of random variables converges to a Gaussian random variable by the central limit theorem.[7] The SINR for the kth user is defined as the ratio of power associated with the kth user's signal over the average power associated with the multiuser interference and noise at the demodulator output. The kth user's performance is then analyzed based on the bit error rate in AWGN with SNR replaced by the SINR for this user. Moreover, if the interference power is much greater than the system noise power, then we can neglect the noise altogether and determine performance based on an AWGN channel analysis with SNR replaced by the signal-to-interference power ratio (SIR) for each user. The SIR for the kth user is defined as the power ratio associated with the kth user's signal over the average power associated with the multiuser interference alone. Multiuser spread-spectrum systems where noise can be neglected in the performance analysis are called *interference limited*, since noise is negligible relative to interference in the performance analysis. For both SINR and SIR, obtaining the average interference power depends on the specific spreading sequences and symbol transmissions of the interfering users, which can be highly complex to analyze. As an alternative, average interference power is often computed assuming random spreading sequences. With this assumption it can be shown that the SIR for a synchronous K-user system with N chips per symbol is given by

$$\text{SIR} = \frac{N}{K-1} \approx \frac{G}{K-1}, \tag{13.45}$$

(see [17, Chap. 2.3]), where $G \approx N$ is the processing gain of the system. Note that this matches the SIR expression (13.6) for arbitrary interference with random signal generation. If noise is taken into account, then the SINR is obtained from (13.45) by adding in noise scaled by the energy per symbol E_s:

$$\text{SINR} = \left(\frac{N_0}{E_s} + \frac{K-1}{G} \right)^{-1}. \tag{13.46}$$

Now consider a more general channel $h_k(t) = \sum_{m=0}^{M-1} \alpha_{km} \delta(t - \tau_{km})$. We assume that $m = 0$ corresponds to the LOS signal path so that, under perfect carrier recovery, we can normalize the delay τ_{k0} and its associated phase shift to zero. The output of the demodulator will again consist of three terms: the first corresponding to the kth

[7] This is true even if the random variables are not i.i.d., as long as they decorrelate.

user's signal, the second corresponding to the interference from other users, and the last an AWGN noise sample, which is not affected by the channel. The signal component associated with the kth user is analyzed the same way as in Section 13.2 for multipath channels: the delayed signal components are attenuated by the autocorrelation of the kth user's spreading code. The multiuser interference is more complicated than before. In particular, assuming the demodulator is synchronized to the LOS component of the kth user, the demodulator output corresponding to the multiuser interference is given by

$$\begin{aligned} I_{kl} &= \frac{2}{T_s}\int_0^{T_s} \sum_{\substack{j=1\\ j\neq k}}^{K}\sum_{m=0}^{M-1} \alpha_{km} s_{j(l-l_{km})} s_{c_j}(i-\tau_{km})\cos(2\pi f_c(\tau-\tau_{km}))s_{c_k}(t)\cos(2\pi f_c t)\,dt \\ &\approx \sum_{\substack{j=1\\ j\neq k}}^{K}\sum_{m=0}^{M-1} \alpha_{km} s_{j(l-l_{km})}\cos(2\pi f_c\tau_{km})\rho_{jk}(\tau_{km}), \end{aligned} \tag{13.47}$$

where $s_{j(l-l_{km})}$ is the symbol associated with the jth user over the $(lT_s-\tau_{km})$th symbol time, which we assume to be constant for simplicity. Comparing (13.44) and (13.47), we see that the multipath channel affects the multiuser interference in two ways. First, there are more interference terms: whereas there were $K-1$ before, we now have $(K-1)M$; so each interfering user contributes M interference terms, one for each multipath component. In addition, the cross-correlation of the codes is no longer taken at delay $\tau = 0$, even though the users are synchronous. In other words, the multipath destroys the synchronicity of the channel. This is significant because orthogonal codes like the Walsh–Hadamard codes typically have zero cross-correlation only at zero delay. So if a Walsh–Hadamard multiuser system operates in a multipath channel, the users will interfere. Equalization of the spreading codes can be used to mitigate this interference.

Example 13.5: Consider a DSSS downlink with bandwidth expansion $N = B_s/B = 100$. Assume the system is interference limited and there is no multipath on any user's channel. Find how many users the system can support under BPSK modulation such that each user has a BER less than 10^{-3}.
Solution: For BPSK we have $P_b = Q(\sqrt{2\gamma_b})$, and $\gamma_b = 6.79\,\text{dB}$ yields $P_b = 10^{-3}$. Since the system is interference limited, we set the SIR equal to the SNR $\gamma_b = 6.79\,\text{dB}$ and solve for K, the number of users:

$$\text{SIR} = \frac{N}{K-1} = \frac{100}{K-1} = 10^{.679} = 4.775. \tag{13.48}$$

Solving for K yields $K \leq 1 + 100/4.77 = 21.96$. Since K must be an integer and we require $P_b \leq 10^{-3}$, it follows that 21.96 must be rounded down to 21 users – although typically a designer would build the system to support 22 users with a slight BER penalty.

13.4.3 Uplink Channels

We now consider DSSS for uplink channels. In multiuser DSSS the spreading code properties are used to separate out the received signals from the different users. The main difference in using DSSS on the uplink versus the downlink is that, in the downlink, both the kth user's signal and the interfering signals from other users pass through the same channel from the transmitter to the kth user's receiver. In an uplink the signals received from each user at the receiver travel through different channels. This gives rise to the *near–far* effect, where users that are close to the uplink receiver can cause a great deal of interference to users farther away, as discussed in more detail below.

The transmitter and channel for each individual user in a K-user uplink are shown in Figure 13.15. As in the downlink, we will focus on in-phase signal analysis. The transmitters are typically not synchronized, since they are not co-located. In general the asynchronous uplink is more complex to analyze than the synchronous uplink and has worse performance. We see from Figure 13.15 that the kth user generates the baseband-modulated signal $x_k(t)$. As in the downlink model, we assume rectangular pulses for $x_k(t)$. The kth user multiplies its baseband signal $x_k(t)$ by its spreading code $s_{c_k}(t)$ and then upconverts to the carrier frequency to form the kth user's transmitted signal $s_k(t)$. (Note that the carrier signals for each user have different phase offsets.) This signal is sent over the kth user's channel, which has impulse response $h_k(t)$. After transmission through their respective channels, all users' signals are summed at the receiver front end together with AWGN $n(t)$.

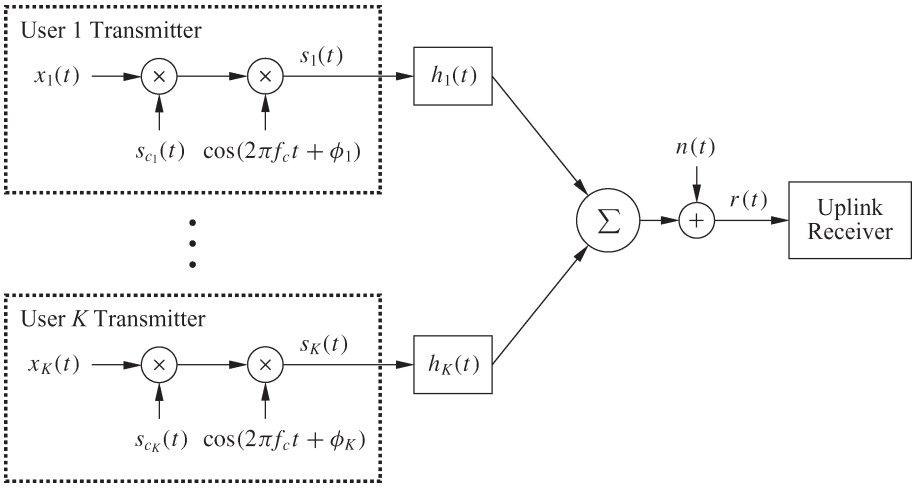


Figure 13.15: DSSS uplink system.

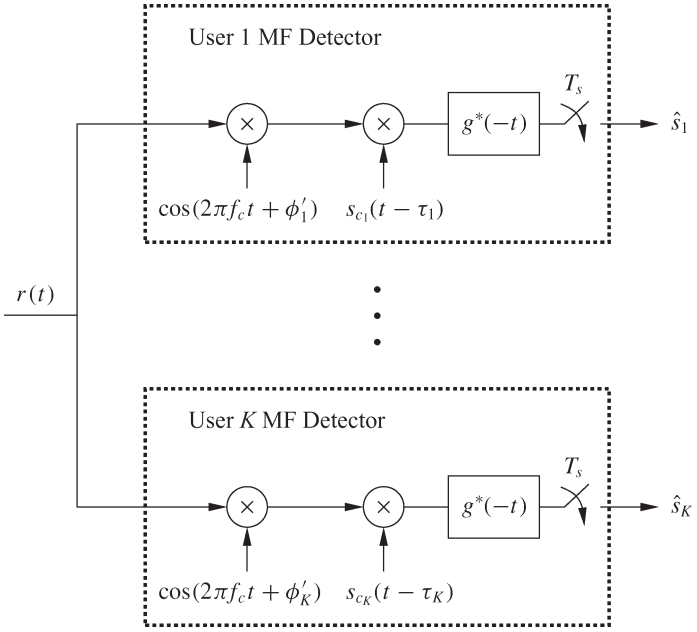


Figure 13.16: Uplink receiver.

The uplink received signal is thus given by

$$r(t) = \left[\sum_{k=1}^{K}(x_k(t)s_{c_k}(t)\cos(2\pi f_c t + \phi_k)) * h_k(t)\right] + n(t). \tag{13.49}$$

The receiver consists of K branches corresponding to the K received signals, as shown in Figure 13.16. We assume (i) that the LOS path of the kth user's channel introduces a delay of τ_k and (ii) that the phase offset resulting from this delay and the transmitter carrier phase offset ϕ_k is matched by the phase offset ϕ'_k obtained by the carrier recovery loop in the kth branch of the receiver. For synchronous users, these delays are all equal. The kth branch downconverts the signal to baseband and then multiplies the received signal by the kth user's spreading code, synchronized to the delay of the kth user's incoming signal. The despread signal is then passed through a matched filter and sampled to obtain an estimate of each user's transmitted symbol over the lth symbol time. Comparing Figures 13.5 and 13.16, we see that the kth branch of the uplink receiver is identical to the matched filter (MF) detector in a single-user DSSS system. Thus, the uplink receiver consists of a bank of K single-user MF detectors, and in the absence of multiuser interference the kth user has identical performance as in a single-user system. With multiuser interference taken into account, the demodulator output of the kth receiver branch over the lth symbol time is given by

$$\begin{aligned}
\hat{s}_k &= \sqrt{\frac{2}{T_s}}\int_0^{T_s}\left[\sum_{j=1}^{K}(x_j(t)s_{c_j}(t)) * h_j^{\mathrm{LP}}(t)\right] s_{c_k}(t-\tau_k)\cos(2\pi f_c t + \phi'_k)\cos(2\pi f_c t + \phi'_j)\,dt \\
&\quad + n_k \\
&= \frac{2}{T_s}\int_0^{T_s}[s_{kl}s_{c_k}(t) * h_k^{\mathrm{LP}}(t)]s_{c_k}(t-\tau_k)\cos^2(2\pi f_c t + \phi'_k)\,dt \\
&\quad + \frac{2}{T_s}\int_0^{T_s}\left[\sum_{\substack{j=1\\ j\neq k}}^{K} s_{ljk}s_{c_j}(t) * h_j^{\mathrm{LP}}(t)\right] s_{c_k}(t-\tau_k)\cos(2\pi f_c t + \phi'_k)\cos(2\pi f_c t + \phi'_j)\,dt \\
&\quad + n_k,
\end{aligned} \tag{13.50}$$

where n_k is the AWGN sample; $h_j^{\mathrm{LP}}(t)$ is the in-phase component of the equivalent lowpass filter for $h_j(t)$, $j = 1, \ldots, K$, with delay normalized to equal zero for the LOS path; and s_{ljk} is the symbol transmitted over the jth user's channel at time $[lT_s - \tau_j + \tau_k, (l+1)T_s - \tau_j + \tau_k]$, which we assume to be constant. If this symbol takes different values on this interval – that is, if it changes values at lT_s then the ISI term is more complicated and involves partial cross-correlations, but the ISI attenuation is roughly the same. Note that (13.50) consists of three separate terms. The first term corresponds to the received signal of the kth user alone, and the last term is the AWGN sample; these two terms are the same as for a single-user system. The second term represents interference from other users in the system, and the interference of the jth user to the kth user ($j \neq k$) depends on the jth user's equivalent lowpass channel $h_j^{\mathrm{LP}}(t)$ and the spreading code properties, as we now show.

Assume that each user's channel introduces only a gain α_j and a delay τ_j, so $h_j^{\mathrm{LP}}(t) = \alpha_j\delta(t - \tau_j)$. Then the demodulator output for the kth branch over the lth symbol time is

$$\hat{s}_k = \alpha_k s_{kl} + I_{kl} + n_l, \tag{13.51}$$

where the first and third terms are the same as for a single-user system with this channel, assuming the spreading code in the receiver is perfectly synchronized to the delay τ_k. Let us now consider the interference term I_{kl}. Substituting $h_j^{\mathrm{LP}}(t) = \alpha_j\delta(t - \tau_j)$ into (13.50) yields

$$
\begin{aligned}
I_{kl} &= \frac{2}{T_s}\int_0^{T_s}\left[\sum_{\substack{j=1\\ j\neq k}}^{K} s_{ljk}s_{c_j}(t)*\alpha_j\delta(t-\tau_j)\right] s_{c_k}(t-\tau_k)\cos(2\pi f_c t+\phi_k')\cos(2\pi f_c t+\phi_j')\,dt \\
&= \frac{1}{T_s}\int_0^{T_s}\left[\sum_{\substack{j=1\\ j\neq k}}^{K} \alpha_j s_{ljk}s_{c_j}(t-\tau_j)\right] s_{c_k}(t-\tau_k)[\cos(\triangle\phi_{kj})+\cos(4\pi f_c t+\phi_k'+\phi_j')]\,dt \\
&\approx \sum_{\substack{j=1\\ j\neq k}}^{K}\alpha_j\cos(\triangle\phi_{kj})s_{ljk}\frac{1}{T_s}\int_0^{T_s} s_{c_j}(t-\tau_j)s_{c_k}(t-\tau_k)\,dt \\
&= \sum_{\substack{j=1\\ j\neq k}}^{K}\alpha_j\cos(\triangle\phi_{kj})s_{ljk}\rho_{jk}(\tau_j-\tau_k),
\end{aligned}
\tag{13.52}
$$

where $\triangle\phi_{kj} = \phi_k' - \phi_j'$ and the approximation is based on $f_c \gg 1/T_s$. We see from (13.52) that, as with the downlink, multiuser interference in the uplink is attenuated by the cross-correlation of the spreading codes. Since typically $\tau_j \neq \tau_k$, users are asynchronous. It follows that orthogonal codes requiring synchronous reception (e.g., Walsh–Hadamard codes) are not typically used on the uplink. Another important aspect of the uplink is that the kth user's symbol and multiuser interference are attenuated by different channel gains. In particular, the kth user's signal is attenuated by the gain α_k, while the interference from the jth user is attenuated by α_j. If $\alpha_j \gg \alpha_k$ then the interference can significantly degrade performance, even though it is reduced by the spreading code cross-correlation.

We now consider interference-limited uplinks. Suppose initially that all users have the same received power. Then the average SINR for asynchronous users on this channel – assuming random spreading codes with N chips per symbol, random start times, and random carrier phases – is given [18] by

$$
\text{SINR} = \left(\frac{K-1}{3N} + \frac{N_0}{E_s}\right)^{-1}. \tag{13.53}
$$

For interference-limited systems we neglect the noise term and obtain the SIR as

$$
\text{SIR} = \frac{3N}{K-1} \approx \frac{3G}{K-1}, \tag{13.54}
$$

where $G \approx N$ is the processing gain of the system. The expressions (13.53) and (13.54) are referred to as the *standard Gaussian approximations* for SINR and SIR. Care must be used in applying these approximations to an arbitrary system, since the SIR and SINR for a given system is heavily dependent on the spreading code properties, timing and carrier phase assumptions, and other characteristics of the system. Modifications to the standard Gaussian approximation have been made to improve its accuracy for practical systems, but these expressions are typically more difficult to work with and don't lead to much greater accuracy than the standard approximations [6, Chap. 9.6]. We can modify (13.54) to approximate the SIR associated with nonrandom spreading codes as

$$
\text{SIR} = \frac{3N}{\xi(K-1)} \approx \frac{3G}{\xi(K-1)}, \tag{13.55}
$$

where ξ is a constant characterizing the code cross-correlation that depends on the spreading code properties and other system assumptions. Under the standard Gaussian assumption we use $\xi = 1$, whereas for PN sequences $\xi = 2$ [19] or $\xi = 3$ [20], depending on the system assumptions.

Suppose now that all $K-1$ interference terms have channel gain $\alpha \gg \alpha_k$. The SIR for the kth user then becomes

$$\mathrm{SIR}(k) = \frac{\alpha_k^2 3N}{\alpha^2 \xi(K-1)} \approx \frac{\alpha_k^2 3G}{\alpha^2 \xi(K-1)} \ll \frac{3G}{\xi(K-1)}, \tag{13.56}$$

so the kth user in the uplink suffers an SIR penalty of α_k^2/α^2 due to the different channel gains. This phenomenon is called the near–far effect, since users far from the uplink receiver will generally have much smaller channel gains to the receiver than those of the interferers. In fading the α_k are random, which typically reduces the code cross-correlation and hence increases the average SIR. The effect of fading can be captured by adjusting ξ in (13.55) to reflect the average cross-correlation under the fading model. The value of ξ then depends on the spreading code properties, the system assumptions, and the fading characteristics [21].

For multipath channels, $h_j^{\mathrm{LP}}(t) = \sum_{m=1}^{M} \alpha_{jm}\delta(t-\tau_{jm})$. Substituting this into (13.52) yields multiuser interference on the kth branch of

$$I_{kl} \approx \sum_{\substack{j=1 \\ j\neq k}}^{K} \sum_{m=1}^{M} \alpha_{jm} \cos(\triangle\phi_{jkm}) s_{ljm} \rho_{jk}(\tau_{jm} - \tau_k), \tag{13.57}$$

where we assume that the kth branch is synchronized to a channel delay of τ_k, that $\triangle\phi_{jkm}$ is the relative phase offset, and that s_{ljm} is the symbol transmitted by the jth user over time $[lT_s - \tau_k + \tau_{jm}, (l+1)T_s - \tau_k + \tau_{jm}]$, which is assumed constant. This interference also contributes to the near-far effect, since any of the multipath components with a large gain relative to the kth user's signal will degrade SIR.

A solution to the near–far effect in DSSS uplink systems is to use power control based on channel inversion, where the kth user transmits signal power P/α_k^2 so that his received signal power is P, regardless of path loss. This will lead to an SIR given by (13.55) for each user. The disadvantage of this form of power control is that channel inversion can require very large transmit power in some fading channels (e.g., infinite power is required in Rayleigh fading). Moreover, channel inversion can cause significant interference to other systems or users operating on the same frequency. In particular, channel inversion can significantly increase the interference between cells in a cellular system. Despite these problems, channel inversion is used on the mobile-to-base station connection in the IS-95 cellular system standard, although the inversion cannot exceed the maximum transmit power constraint.

Example 13.6: Consider a DSSS uplink system with processing gain $G = B_s/B = 100$. Assume the system is interference limited and there is no multipath on any user's channel. Suppose user k has a received power that is 6 dB less than the other users. Find the number of users that the system can support under BPSK modulation such that each user has a BER less than 10^{-3}. Make the computation both for random codes under the standard Gaussian assumption, $\xi = 1$, and for PN codes with $\xi = 3$.

Solution: As in the previous example, we require $\mathrm{SIR} = \gamma_b = 6.79\,\mathrm{dB} = 4.775$ for $P_b = 10^{-3}$. We again set the SIR equal to the SNR $\gamma_b = 4.775$ and solve for K to find the maximum number of users that the system can support. Since this is an asynchronous system with $\alpha_k^2/\alpha^2 = .251(-6\,\mathrm{dB})$, we have

$$\mathrm{SIR} = \frac{\alpha_k^2 3N}{\alpha^2 \xi(K-1)} = \frac{.251(300)}{\xi(K-1)} = \frac{75.3}{\xi(K-1)} = 4.775. \tag{13.58}$$

Solving for K yields $K \leq 1 + 75.3/4.77\xi = 16.78$ for $\xi = 1$ and $K \leq 6.26$ for $\xi = 3$, so the system can only support between 6 and 16 users – up to a factor of 3 less than in the prior example for the downlink, which is due to

the asynchronicity of the uplink and the near–far effect. This example also illustrates the sensitivity of the system capacity calculation to assumptions about the spreading code properties as captured by ξ. Since the SIR is roughly proportional to $1/(\xi K)$, the number of users the system can support for a given SIR is roughly proportional to $1/\xi$.

13.4.4 Multiuser Detection

Interference signals in SSMA need not be treated as noise. If the spreading code of the interference signal is known, then this knowledge can be used to mitigate the effects of the multiple access interference (MAI). In particular, if all users are detected simultaneously, then interference between users can be subtracted out, which either improves performance or, for a given performance, allows more users to share the channel. Moreover, if all users are detected simultaneously then the near–far effect can aid in detection, since users with strong channel gains are more easily detected (for subsequent cancellation) than if all users had the same channel gains. A DSSS receiver that exploits the structure of multiuser interference in signal detection is called a multiuser detector (MUD). Multiuser detectors are not typically used on downlink channels for several reasons. First, downlink channels are typically synchronous, so they can eliminate all interference by using orthogonal codes as long as the channel doesn't corrupt code orthogonality. Moreover, the kth user's receiver in a downlink is typically limited in terms of power and/or complexity, which makes it difficult to add complex MUD functionality. Finally, the uplink receiver must detect the signals from all users anyway, so any receiver in the uplink is by definition a multiuser detector, albeit not necessarily a good one. By contrast, the kth user's receiver in the downlink need only detect the signal associated with the kth user. For these reasons, work on MUD has primarily focused on DSSS uplink systems, and that is the focus of this section.

Multiuser detection was pioneered by Verdú in [22, 23], where the optimum joint detector for the DSSS asynchronous uplink channel was derived. This derivation assumes an AWGN channel with different channel gains for each user. The optimum detector for this channel chooses the symbol sequences associated with all K users that minimize the MSE between the received signal and the signal that would be generated by these symbol sequences. Because the channel is asynchronous, the entire received waveform must be processed for optimal detection over any one symbol period. The reason is that symbols from other users are not aligned in time, hence all symbols that overlap in the given interval of interest must be considered; by applying to same reasoning to the overlapping symbols, we see that it is not possible to process the signal over any finite interval and still preserve optimality. The optimal MUD for the asynchronous case was shown in [23] to consist of a bank of K single-user matched filter detectors, followed by a Viterbi sequence detection algorithm to jointly detect all users. The Viterbi algorithm has 2^{K-1} states and complexity that grows as 2^K, assuming binary modulation.

For synchronous users, the optimal detection becomes simpler: only one symbol interval needs to be considered in the optimal joint detection, so sequence detection is not needed. In particular, the complexity of this detector grows as $2^K/K$, in contrast to the 2^K growth associated with sequence detection. Consider a two-user synchronous uplink with gain α_k on channel k and binary modulation. The equivalent lowpass received signal over one bit time is

$$r(t) = \alpha_1 \mathrm{b}_1 s_{c_1}(t) + \alpha_2 \mathrm{b}_2 s_{c_2}(t) + n(t), \tag{13.59}$$

where b_k is the bit transmitted by the kth user over the given bit time. The optimum (maximum likelihood) detector outputs the pair $\mathbf{b}^* = (\mathrm{b}_1^*, \mathrm{b}_2^*)$ that satisfies

$$\mathbf{b}^* = \arg \min_{(\mathrm{b}_1, \mathrm{b}_2)} \left[-\frac{1}{2\sigma^2} \int_0^{T_s} [r(t) - \alpha_1 \mathrm{b}_1 s_{c_1}(t) - \alpha_2 \mathrm{b}_2 s_{c_2}(t)]^2 \, dt \right], \tag{13.60}$$

where σ^2 is the noise power. This is equivalent to finding $(\mathrm{b}_1^*, \mathrm{b}_2^*)$ to maximize the cost function

$$L(\mathrm{b}_1, \mathrm{b}_2) = \alpha_1 \mathrm{b}_1 r_1 + \alpha_2 \mathrm{b}_2 r_2 - \alpha_1 \alpha_2 \mathrm{b}_1 \mathrm{b}_2 \rho_{12}, \tag{13.61}$$

where

$$r_k = \int_0^{T_b} r(t) s_{c_k}(t)\, dt \tag{13.62}$$

and

$$\rho_{jk} = \int_0^{T_b} s_{c_k}(t) s_{c_j}(t)\, dt. \tag{13.63}$$

This analysis easily extends to K synchronous users. Here we can express $\mathbf{r} = (r_1, \ldots, r_K)^T$ in matrix form as

$$\mathbf{r} = \mathbf{RAb} + \mathbf{n} \tag{13.64}$$

(see [24]), where $\mathbf{b} = (\mathrm{b}_1, \ldots, \mathrm{b}_K)^T$ is the bit vector associated with the K users over the given bit time, $\mathbf{A}$ is a diagonal $K \times K$ matrix of the channel gains α_k, and $\mathbf{R}$ is a $K \times K$ matrix of the cross-correlations between the spreading codes. The optimal choice of bit sequence $\mathbf{b}^*$ is obtained by choosing the sequence to maximize the cost function

$$L(\mathbf{b}) = 2\mathbf{b}^T \mathbf{Ar} - \mathbf{b}^T \mathbf{ARAb}. \tag{13.65}$$

Unfortunately, maximizing (13.65) for K users also has complexity that grows as 2^K (the same as in the asynchronous case), assuming a search tree is used for the optimization. In addition to the high complexity of the optimal detector, it has the drawback of requiring knowledge of the channel amplitudes α_k.

The complexity of MUD can be decreased at the expense of optimality. Many suboptimal MUDs have been developed with various trade-offs with respect to performance, complexity, and requirements regarding channel knowledge. Suboptimal MUDs fall into two broad categories: linear and nonlinear. Linear MUDs apply a linear operator or filter to the output of the matched filter bank in Figure 13.16. These detectors have complexity that is linear in the number of users, a significant complexity improvement over the optimal detector. The most common linear MUDs are the decorrelating detector [25] and the minimum mean-square error (MMSE) detector. The decorrelating detector simply inverts the matrix $\mathbf{R}$ of cross-correlations, resulting in

$$\hat{\mathbf{b}}^* = \mathbf{R}^{-1}\mathbf{r} = \mathbf{R}^{-1}[\mathbf{RAb} + \mathbf{n}] = \mathbf{Ab} + \mathbf{R}^{-1}\mathbf{n}. \tag{13.66}$$

The inverse exists for most cases of interest. In the absence of noise, the resulting bit sequence equals the original sequence, scaled by the channel gains. In addition to its simplicity, this detector has other appealing features: it completely removes MAI, and it does not require knowledge of the channel gains. However, the decorrelating detector can lead to noise enhancement, since the noise vector is multiplied by the matrix inverse. Thus, decorrelating MUD is somewhat analogous to zero-forcing equalization as described in Section 11.4.1: all MAI can be removed, but at the expense of noise enhancement.

The MMSE detector finds the matrix $\mathbf{D}$ such that multiplication of the filter bank output by $\mathbf{D}$ minimizes the expected MSE between $\mathbf{D}$ and the transmitted bit sequence $\mathbf{b}$. In other words, the matrix $\mathbf{D}$ satisfies

$$\arg\min_{\mathbf{D}} \mathbf{E}[(\mathbf{b} - \mathbf{Dr})^T (\mathbf{b} - \mathbf{Dr})]. \tag{13.67}$$

The optimizing $\mathbf{D}$ is given [17, 24, 26] by

$$\mathbf{D} = (\mathbf{R} + .5N_0\mathbf{I})^{-1}. \tag{13.68}$$

Note that, in the absence of noise, the MMSE detector is the same as the decorrelating detector. However, it has better performance at low SNRs because it balances removal of the MAI with noise enhancement. This is analogous to the MMSE equalizer design for ISI channels, described in Section 11.4.2.

Nonlinear MUDs can have much better performance than linear detectors – although not necessarily in all cases, especially with little or no coding [27]. The most common nonlinear MUD techniques are multistage detection, decision-feedback detection, and successive interference cancellation. In a multistage detector, each stage consists of the conventional matched filter bank. The nth stage of the detector uses decisions of the $(n-1)$th stage to cancel the MAI at its input. The multistage detector can be applied to either synchronous [28] or asynchronous [29] systems. The decision-feedback detector is based on the same premise as a decision-feedback equalizer. It consists of a feedforward and feedback filter, where the feedforward filter is the Cholesky factorization of the correlation matrix $\mathbf{R}$. The decision-feedback MUD can be designed for either synchronous [30] or asynchronous [31] systems. These detectors require knowledge of the channel gains and can also suffer from error propagation when decision errors are fed back through the feedback filter. In interference cancellation, an estimate of one or more users is made and then the MAI caused to other users is subtracted out [32]. Interference cancellation can be done in parallel, where all users are detected simultaneously and then cancelled out [33, 34], or sequentially, where users are detected one at a time and then subtracted out from users yet to be detected [35]. Parallel cancellation has a lower latency and is more robust to decision errors. However, its performance suffers owing to the near–far effect, when some users have much weaker received powers than others. Under this unequal power scenario, successive interference cancellation can outperform parallel cancellation [24]. In fact, successive cancellation theoretically achieves Shannon capacity of the uplink channel, as will be discussed in Chapter 14, and in practice approaches Shannon capacity [36]. Successive interference cancellation suffers from error propagation, which can significantly degrade performance, but this degradation can be partially offset through power control [37].

A comprehensive treatment of different MUDs and their performance can be found in [17], and shorter tutorials are provided in [24, 32]. Combined equalization and MUD is treated in [38]. Multiuser detection for multirate CDMA, where different users have different data rates, is analyzed in [39]. Blind, space-time, and turbo multiuser detectors are developed in [40]. Spectral efficiencies of the different detectors have been analyzed in [27].

13.4.5 Multicarrier CDMA

Multicarrier CDMA (MC-CDMA) is a technique that combines the advantages of OFDM and CDMA. It is very effective both at combating ISI and as a mechanism for allowing multiple users to share the same channel. The basic block diagram for a baseband single-user multicarrier CDMA system is shown in Figure 13.17. The data symbol s_l is sent over all N subchannels. On the ith subchannel, s_l is multiplied by the ith chip c_j of a spreading sequence $s_c(t)$, where $c_j = \pm 1$. This is similar to the standard spread-spectrum technique, except that multiplication with the spreading sequence is done in the frequency domain rather than in the time domain. The frequency spread data $(s_l c_1, s_l c_2, \ldots, s_l c_N)$ is then multicarrier modulated in the standard manner: the parallel sequence is passed through an inverse fast Fourier transform (IFFT) and a parallel-to-serial converter and then is D/A converted to produce the modulated signal $s(t)$, where $S(f)$ is as shown in Figure 13.17 for subchannel carrier frequencies $(f_1, \ldots, f_N)$.

Assume the MC-CDMA signal is transmitted through a frequency-selective channel with a constant channel gain of α_i on the ith subchannel and with AWGN $n(t)$. The receiver performs the reverse operations of the transmitter – passing the received signal through an A/D converter, a serial-to-parallel converter, and an FFT – to recover the symbol transmitted over the ith subchannel. The subchannel symbol received on the ith subchannel is multiplied by the ith chip c_i and a weighting factor β_i; these terms are then summed together for the final symbol

estimate $\hat{s}_l$.

In a multiuser MC-CDMA system, each user k modulates his signal as in Figure 13.17 but with a unique spreading code $s_{c_k}(t)$. So, for a two-user system, user 1 would use the spreading code $s_{c_1}(t)$ with chips $(c_1^1, \ldots, c_N^1)$ resulting in a transmitted signal $s_1(t)$, and user 2 would use the spreading code $s_{c_2}(t)$ with chips $(c_1^2, \ldots, c_N^2)$ resulting in a transmitted signal $s_2(t)$. If the users transmit simultaneously then their signals are added "in the air" as shown in Figure 13.18, where s_l^1 (resp., s_l^2) is the symbol corresponding to user 1 (resp., user 2) over the lth symbol time. The interference between users in this system is reduced by the cross-correlation of the spreading codes, as in standard spread spectrum without multicarrier. However, each user benefits from the frequency diversity of spreading its signal over independently fading subchannels. This typically leads to better performance than in standard spread spectrum.

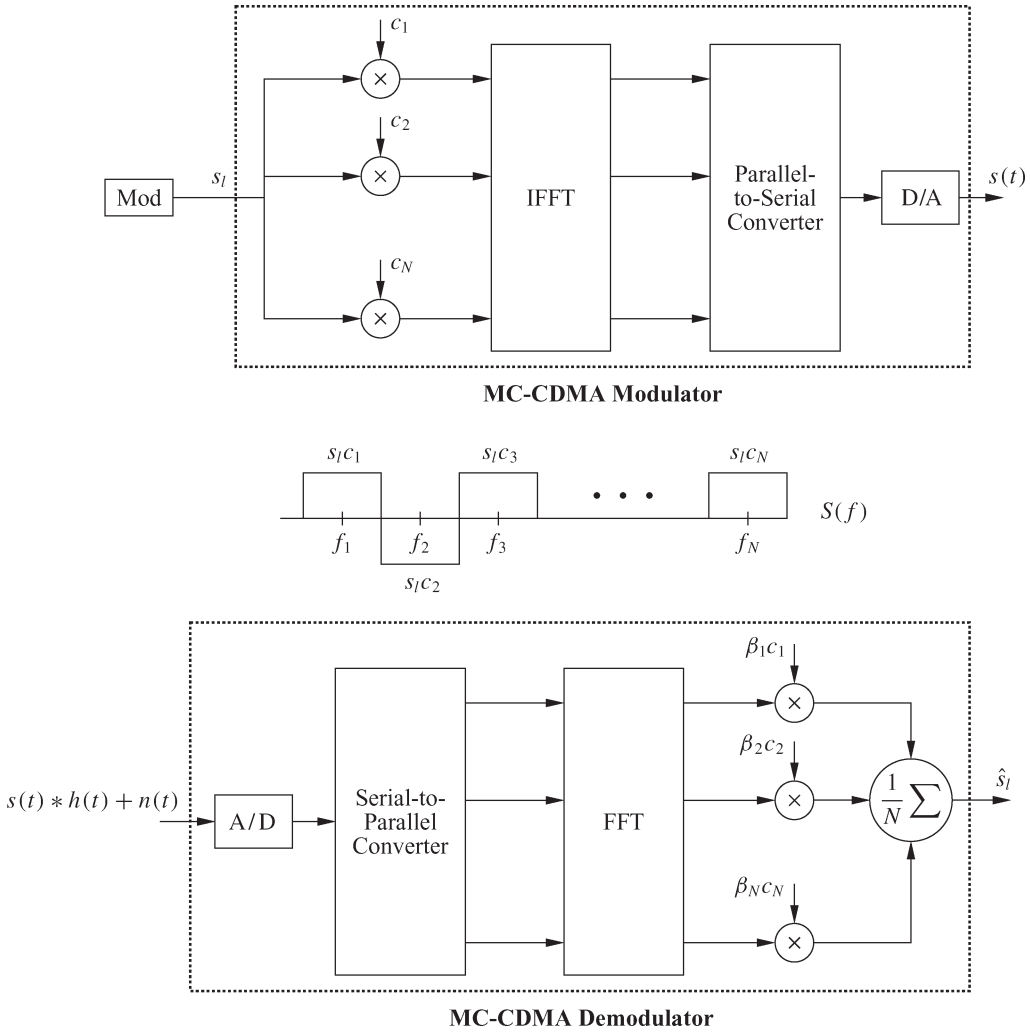


Figure 13.17: Multicarrier CDMA system.

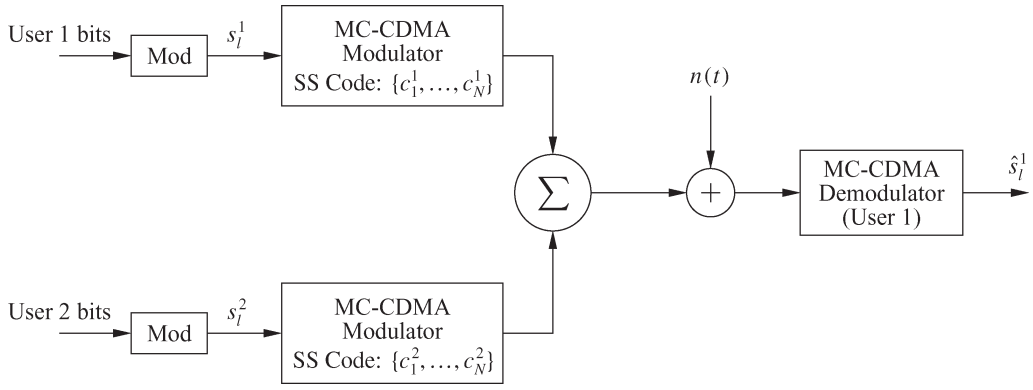


Figure 13.18: Two-user multicarrier CDMA system.

13.5 Multiuser FHSS Systems

Multiuser FHSS is accomplished by assigning each user a unique spreading code sequence $s_{c_i}(t)$ to generate its hop pattern. If the spreading codes are orthogonal and if the users are synchronized in time, then the different users never collide, and performance of each user is the same as in a single-user FH system. However, if the users are asynchronous or if nonorthogonal codes are used, then multiple users will collide by occupying a given channel simultaneously. The symbols transmitted during that time are likely to be in error, so multiuser FHSS typically uses error correction coding to compensate for collisions.

Multiuser FHSS, also referred to as FH-CDMA or FH-SSMA, is mainly applied to uplink channels. This access method is the preferred method for military applications because of the protection from jamming and low probability of interception and detection characteristic of FH systems. FH-CDMA was also proposed as a candidate for second-generation digital cellular systems [2, Chap. 9.4] but was not adopted. Trade-offs between frequency hopping and direct sequence for multiple access are discussed in [8, Chap. 11; 41; 42]. In fact, most analyses indicate that FH-CDMA is inferior to DS-CDMA as a multiple access method in terms of the number of users that can share the channel simultaneously, at least under asynchronous operation. In addition, FH-CDMA systems typically cause interference at much larger distances than DS-CDMA systems, since the interference isn't mitigated by bandwidth spreading [8, Chap. 11]. The main benefit of FH over DS is its robustness to the near-far effect: avoiding interference by hopping – rather than suppressing it by the processing gain – mitigates the performance degradation caused by strong interferers. Hybrid techniques using FH along with another multiple access method have been used to exploit the benefits of both. For example, the GSM digital cellular standard uses a combination of time division and slow frequency hopping, where the frequency hopping is used primarily to average out interference from other cells. FH-CDMA is also used in the Bluetooth system. Bluetooth operates in the unlicensed 2.4-GHz band, and FH was chosen because it can be used with noncoherent FSK modulation, a low-cost and energy-efficient modulation technique.

Chapter 13 Problems

1. In this problem we derive the SIR ratio (13.6) for a randomly spread signal with interference. The correlator output of the ith receiver branch in this system is given by (13.5) as

$$x_i = \int_0^T x(t)s_i(t)\,dt = \sum_{j=1}^{N}(s_{ij}^2 + I_j s_{ij}).$$

   (a) Show that the conditional expectation of x_i, conditioned on the transmitted signal $s_i(t)$, is $\mathbf{E}[x_i|s_i(t)] = E_s$.

   (b) Assuming equiprobable signaling ($p(s_i(t)) = 1/M$), show that $\mathbf{E}[x_i] = E_s/M$.

   (c) Show that Var $[x_i|s_i(t)] = E_s E_J/N$.

   (d) Show that (again with equiprobable signaling) Var $[x_i] = E_s E_J/NM$.

   (e) The SIR is given by

$$\text{SIR} = \frac{\mathbf{E}[x_i]^2}{\text{Var}[x_i]}.$$

   Show that

$$\text{SIR} = \frac{E_s}{E_j} \cdot \frac{N}{M}.$$

2. Sketch the transmitted DSSS signal $s(t)s_c(t)$ over two bit times $[0, 2T_b]$ assuming that $s(t)$ is BPSK modulated with carrier frequency 100 MHz and $T_s = 1$ μs. Assume the first data bit is a 1 and the second data bit is a 0. Assume also that there are ten chips per bit and that the chips alternate between ± 1, with the first chip equal to $+1$.

3. Consider an FH system transmitted over a two-ray channel, where the reflected path has delay $\tau = 10$ μs relative to the LOS path. Assume the receiver is synchronized to the hopping of the LOS path.

   (a) For what hopping rates will the system exhibit no fading?

   (b) Assume an FFH system with hop time $T_c = 50$ μs and symbol time $T_s = .5$ ms. Will this system exhibit no fading, flat fading, or frequency-selective fading?

   (c) Assume a SFH system with hop time $T_c = 50$ μs and symbol time $T_s = .5$ μs. Will this system exhibit no fading, flat fading, or frequency-selective fading?

4. In this problem we explore the statistics of the DSSS receiver noise after it is multiplied by the spreading sequence. Let $n(t)$ be a random noise process with autocorrelation function $\rho_n(\tau)$ and let $s_c(t)$ be a zero-mean random spreading code, independent of $n(t)$, with autocorrelation $\rho_c(\tau)$. Let $n'(t) = n(t)s_c(t)$.

   (a) Find the autocorrelation and power spectral density of $n'(t)$.

   (b) Show that if $\rho_c(\tau) = \delta(\tau)$ then $n'(t)$ is zero mean with autocorrelation function $\rho_n(\tau)$ –that is, it has the same statistics as $n(t)$, so the statistics of $n(t)$ are not affected by its multiplication with $s_c(t)$.

   (c) Find the autocorrelation $\rho_{n'}(\tau)$ of $n'(t)$ if $n(t)$ is zero-mean AWGN and $s_c(t)$ is a maximal linear code with autocorrelation given by (13.19). What happens to $\rho_{n'}(\tau)$ as $N \to \infty$ in (13.19)?

5. Show that, for any real periodic spreading code $s_c(t)$, its autocorrelation $\rho_c(\tau)$ over one period is symmetric about τ and reaches its maximum value at $\tau = 0$.

6. Show that, if $s_c(t)$ is periodic with period T, then the autocorrelation for time-shifted versions of the spreading code depends only on the difference of their time shifts:

$$\frac{1}{T}\int_0^T s_c(t-\tau_0)s_c(t-\tau_1)\,dt = \rho_c(\tau_1-\tau_0).$$

7. Show that, for any periodic spreading code $s_c(t)$ with period T, its autocorrelation $\rho_c(t)$ is periodic with the same period.

8. Show that the power spectral density of $s_c(t)$ for a maximal linear spreading code with period $NT_c = T_s$ is given by

$$P_{s_c}(f) = \sum_{m=-\infty}^{\infty} \frac{N+1}{N^2}\mathrm{sinc}^2\left(\frac{m}{N}\right)\delta\left(f-\frac{m}{T_s}\right).$$

   Also plot this spectrum for $N = 100$ and $T_s = 1$ μs over the frequency range $-10/T_s \le f \le 10/T_s$.

9. Show that m-sequences as well as random binary spreading sequences have the balanced, run-length, and shift properties.

10. Suppose that an unmodulated carrier $s(t)$ is spread using a maximal linear code $s_c(t)$ with period T and then is transmitted over a channel with an impulse response of $h(t) = \alpha_0\delta(t-\tau_0) + \alpha_1\delta(t-\tau_1)$. The corresponding received signal $r(t)$ is input to the synchronization loop shown in Figure 13.9. Find the function $w(\tau)$ output from the integrator in this loop as a function of τ. What will determine which of these two multipath components the control loop for coarse acquisition locks to?

11. Find the outage probability relative to $P_b = 10^{-6}$ for a three-branch RAKE receiver with DPSK signal modulation, independent Rayleigh fading on each branch, and a branch SNR/bit (prior to despreading) of 10 dB. Assume the code autocorrelation associated with maximal linear codes with $K = N = 2^n - 1 = 15$. Assume also that the code in the first branch is perfectly aligned, but that the code in the second branch is offset by $T_c/4$ and the code in the third branch is offset by $T_c/3$. Assume selection combining diversity and neglect the interference due to other multipath components in your SNR calculations.

12. Consider a spread-spectrum signal transmitted over a multipath channel with a LOS component and a single multipath component, where the delay of the multipath relative to the LOS is greater than the chip time T_c. Consider a two-branch RAKE receiver, with one branch corresponding to the LOS component and the other to the multipath component. Assume that, with perfect synchronization in both branches, the incoming signal component at each branch (after despreading) has power that is uniformly distributed between 6 mW and 12 mW. The total noise power in the despread signal bandwidth is 1 mW. Suppose, however, that only the first branch is perfectly synchronized, while the second branch has a timing offset of $T_c/2.366$. The code autocorrelation is that of a maximal linear code with $N \gg 1$. The two branches of the RAKE are combined using maximal-ratio combining with knowledge of the timing offset.

    (a) What is the average SNR at the combiner output?

    (b) What is the distribution of the combiner output SNR?

(c) What is the outage probability for DPSK modulation with a BER of 10^{-4}?

13. This problem illustrates the benefits of RAKE receivers and the optimal choice of multipath components for combining when the receiver complexity is limited. Consider a multipath channel with impulse response

$$h(t) = \alpha_0\delta(t) + \alpha_1\delta(t-\tau_1) + \alpha_2\delta(t-\tau_2).$$

The α_i are Rayleigh fading coefficients, but their expected power varies (because of shadowing) such that $\mathbf{E}[\alpha_0^2] = 5$ with probability .5 and 10 with probability .5, $\mathbf{E}[\alpha_1^2] = 0$ with probability .5 and 20 with probability .5, and $\mathbf{E}[\alpha_2^2] = 5$ with probability .75 and 10 with probability.25 (all units are linear). The transmit power and noise power are such that a spread-spectrum receiver locked to the ith multipath component will have an SNR of α_i^2 in the absence of the other multipath components.

(a) Assuming maximal linear codes, a bit time T_b, and a spread-spectrum receiver locked to the LOS signal component (with zero delay and gain α_0), find the values of τ_1 and $\tau_2 (0 \leq \tau_1 \leq \tau_2 < T_b)$ for which their corresponding multipath components will be attenuated by $-1/N$, where N is the number of chips per bit.

For the rest of this problem, assume spreading codes with autocorrelation equal to a delta function.

(b) What is the outage probability of DPSK modulation at an instantaneous $P_b = 10^{-3}$ for a single-branch spread-spectrum receiver locked to the LOS path?

(c) Find the outage probability of DPSK modulation at an instantaneous $P_b = 10^{-3}$ for a three-branch RAKE receiver, where each branch is locked to one of the multipath components and SC is used to combine the paths.

(d) Suppose receiver complexity is limited such that only a two-branch RAKE with SC can be built. Find which two multipath components the RAKE should lock to in order to minimize the outage probability of DPSK modulation at $P_b = 10^{-3}$, and then find this minimum outage probability.

14. This problem investigates the performance of a RAKE receiver when the multipath delays are random. Consider a DS spreading code with chip time T_c and autocorrelation function

$$\rho_c(t) = \begin{cases} 1 & -T_c/2 < t < T_c/2, \\ 0 & \text{else.} \end{cases}$$

Suppose this spreading code is used to modulate a DPSK signal with bit time $T_b = 10T_c$. The spread signal is transmitted over a multipath channel, where the channel is modeled using the discrete-time tapped delay model described in Section 3.4, with a tap separation of T_c and a total multipath spread $T_m = 5T_c$. Thus, the model has five multipath "bins", where the ith bin has at most one multipath component of delay $(i - .5)T_c$. The distribution of the multipath component in each bin is independent of the components in all the other bins. The probability of observing a multipath component in bin i is.75 and, conditioned on having a multipath component in bin i, the amplitude of the ith multipath component after despreading is Rayleigh distributed with average SNR/bit of $S_i = 20/i$ for $i = 1, 2, \ldots, 5$ (in linear units). Thus, the average power is decreasing relative to the distance that the multipath component has traveled.

At the receiving end is a five-branch SC-diversity RAKE receiver with each branch synchronized to one of the multipath bins. Assuming a target BER of 10^{-3}, compute the outage probability of the RAKE receiver

output. Compare this with the outage probability for the same BER if there is (with probability 1) a multipath component in each bin and if each multipath component after despreading has an average SNR/bit of $S_i = 20$ (linear units).

15. Direct-sequence spread-spectrum signals are often used to make channel measurements, since the wideband spread-spectrum signal has good resolution of the individual multipath components in the channel. Channel measurement with spread spectrum, also called *channel sounding*, is performed using a receiver with multiple branches synchronized to the different chip delays. Specifically, an unmodulated spreading sequence $s_c(t)$ is sent through the channel $h(t)$, as shown in Figure 13.19. The receiver has $N + 1$ branches synchronized to different chip delays. The β_i output from the ith branch approximates the channel gain associated with that delay, so that the approximate channel model obtained by the channel sounding shown in the figure is $\hat{h}(t) = \sum_{i=0}^{N} \beta_i \delta(t - iT_c)$.

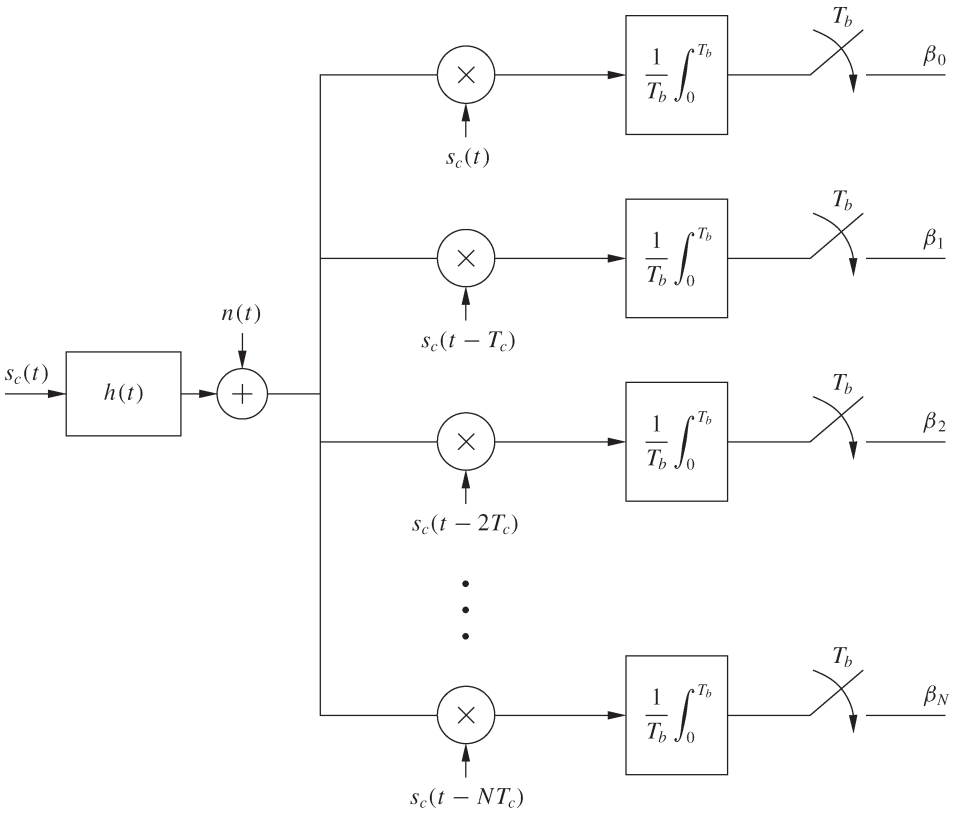


Figure 13.19: Channel sounder for Problem 13-15.

Assume that the autocorrelation function for $s_c(t)$ is

$$\rho_c(\tau) = \frac{1}{T_b}\int_0^{T_b} s_c(t)s_c(t-\tau)\,dt = \begin{cases} 1 - |\tau|/T_c & |\tau| < T_c, \\ 0 & |\tau| \geq T_c. \end{cases}$$

(a) Show that if $h(t) = \sum_{i=0}^{N} \alpha_i \delta(t - iT_c)$ then, in the absence of noise ($n(t) = 0$), the channel sounder of Figure 13.19 will output $\beta_i = \alpha_i$ for all i.

(b) Again neglecting noise, if $h(t) = a\delta(t) + b\delta(t - 1.2T_c) + c\delta(t - 3.5T_c)$, what approximation $\hat{h}(t)$ will be obtained by the channel sounder?

(c) Now assume that the channel sounder yields a perfect estimate of the channel $\hat{h}(t) = h(t) = \beta_0\delta(t) + \beta_1\delta(t - T_c) + \beta_2\delta(t - 2T_c)$, where the β_i are all independent Rayleigh fading random variables. Consider a three-branch RAKE receiver with the ith branch perfectly synchronized to the ith multipath component of $h(t)$ and with an average SNR/bit on each branch (after despreading) of 10 dB. Find P_{out} for DPSK modulation with a target BER of 10^{-3} under maximal-ratio combining in the RAKE. Do the same calculation for selection combining in the RAKE.

16. Find the values of the autocorrelation and cross-correlation for Gold codes, Kasami codes from the small set, and Kasami codes from the large set for $n = 8$. Also, find the number of such Kasami codes for both the small set and the large set.

17. Find the Hadamard matrix for $N = 4$ and show that the spreading codes generated by the rows of this matrix are orthogonal, assuming synchronous users i.e., show $\rho_{ij}(0) = 0 \forall i \neq j)$. Also find the cross-correlation between all pairs of users assuming a timing offset of $T_c/2$ between users; that is, find

$$\rho_{ij}\left(\frac{T_c}{2}\right) = \frac{1}{T_s}\int_0^{T_s} s_{c_i}(t)s_{c_j}\left(t - \frac{T_c}{2}\right)dt = \frac{1}{N}\sum_{n=1}^{N} s_{c_i}(nT_c)s_{c_j}\left(nT_c - \frac{T_c}{2}\right)$$

for all pairs of codes.

18. Consider an asynchronous DSSS multiple access channel (MAC) system with bandwidth expansion $N = B_s/B = 100$ and $K = 40$ users. Assume the system is interference limited and there is no multipath on any user's channel. Find the probability of error for user k under BPSK modulation, assuming random codes with the standard Gaussian assumption and assuming this user is in a deep fade, with received power that is 6 dB less than the other users. Would this change if the users could be synchronized?

19. Show that the vector $\mathbf{r} = (r_1, \ldots, r_K)^T$ for r_k given by (13.62) can be expressed by the matrix equation (13.64). What are the statistics of $\mathbf{n}$ in this expression?

20. Show that the maximum likelihood detector for a K-user synchronous MAC receiver chooses the vector $\mathbf{b}$ to maximize the cost function given by (13.65).

21. Here we illustrate the use of multiple spreading codes in single-user CDMA systems for adaptive modulation or diversity gain. The BER for user k in a K-user DS-CDMA system, where each user transmits his BPSK modulated bit sequence at a rate R bps along his spreading code, is given by:

$$\text{BER}_k = Q\left(\sqrt{\frac{2P_k(\gamma_k)\gamma_k}{\frac{1}{N}\sum_{i=1,i\neq k}^{K} P_i(\gamma_i)\gamma_i + 1}}\right), \tag{13.69}$$

where N is the spreading factor (processing gain), γ_i is the ith user's channel power gain, and $P_i(\gamma_i)$ is i's transmit power when his channel gain is γ_i. Note that noise power has been normalized to unity and that the receiver demodulates each spreading sequence, treating other sequences as noise (matched filter detector). The system has a *single* user who can simultaneously transmit *up to two* spreading sequences, modulating each with an independent BPSK bit stream at rate R bps (on each stream).

(a) Assume that the user's channel fade is γ, and assume that he splits his total transmit power $P(\gamma)$ equally among the transmitted sequences. Note that the user has three options: he can transmit nothing, one BPSK modulated spreading sequence, or both spreading sequences BPSK modulated with independent

bits. Based on the BER expression for the multiuser case (13.69), explain why the BER for the single-user multirate DS-CDMA system is given by

$$\text{BER} = Q\left(\sqrt{2P(\gamma)\gamma}\right) \tag{13.70}$$

if the user transmits only one spreading sequence and by

$$\text{BER} = Q\left(\sqrt{\frac{P(\gamma)\gamma}{P(\gamma)\gamma/2N+1}}\right) \tag{13.71}$$

when he transmits both spreading sequences together. What are the rates achieved in both of these cases?

(b) Assume the channel is known perfectly to the transmitter as well as the receiver, and assume that γ is distributed according to the distribution $p(\gamma)$. We want to develop an adaptive rate and power strategy for this channel. Since we do not use error correction coding, the user needs to keep his BER below a threshold P_b^0 for all transmitted bits. Assume an average transmit power constraint of unity: $\int_0^\infty P(\gamma)p(\gamma)\ d\gamma = 1$. We have a finite discrete set of possible rates and so, as with narrowband adaptive modulation, the optimal adaptive rate policy is to send no data when γ is below a cutoff threshold γ_0, one data stream when $\gamma_0 \leq \gamma < \gamma_1$, and both data streams when $\gamma > \gamma_1$. Find the power adaptation strategy that exactly meets the BER target P_b^0 for this adaptive rate strategy as a function of the thresholds γ_0 and γ_1.

(c) Given the adaptive rate and power strategy obtained in part (b), solve the Lagrangian optimization to find γ_0 and γ_1 as a function of the Lagrangian λ.

22. You work for a company that wants to design a next-generation cellular system for voice plus high-speed data. The FCC has decided to allocate 100 MHz of spectrum for this system based on whatever standard is agreed to by the various industry players. You have been charged with designing the system and pushing your design through the standards body. You should describe your design in as much detail as possible, paying particular attention to how it will combat the impact of fading and ISI and to its capacity for accommodating both voice and data. Also develop arguments explaining why your design is better than competing strategies.

Bibliography

[1] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications–A tutorial," *IEEE Trans. Commun*., pp. 855–84, May 1982.

[2] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[3] V. V. Veeravalli and A. Mantravadi, "The coding–spreading trade-off in CDMA systems," *IEEE J. Sel. Areas Commun*., pp. 396–408, February 2002.

[4] S. W. Golomb, *Shift Register Sequences*, Holden-Day, San Francisco, 1967.

[5] A. J. Viterbi, *CDMA Principles of Spread Spectrum Communications*, Addison-Wesley, Reading, MA, 1995.

[6] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[7] O. C. Mauss, F. Classen, and H. Meyr, "Carrier frequency recovery for a fully digital direct-sequence spread-spectrum receiver: A comparison," *Proc. IEEE Veh. Tech. Conf*., pp. 392–5, May 1993.

[8] R. C. Dixon, *Spread Spectrum Systems with Commercial Applications*, 3rd ed., Wiley, New York, 1994.

[9] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*, McGraw-Hill, New York, 1994.

[10] B. Sklar, *Digital Communications–Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[11] G. L. Turin, "Introduction to spread spectrum antimultipath techniques and their application to urban digital radio," *Proc. IEEE*, pp. 328–53, March 1980.

[12] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*: *A Unified Approach to Performance Analysis*, Wiley, New York, 2000.

[13] R. Gold, "Optimum binary sequences for spread-spectrum multiplexing," *IEEE Trans. Inform. Theory*, pp. 619–21, October 1967.

[14] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Commun. Mag*., pp. 48–54, September 1998.

[15] N. B. Mehta, L. J. Greenstein, T. M. Willis, and Z. Kostic, "Analysis and results for the orthogonality factor in WCDMA downlinks," *IEEE Trans. Wireless Commun*., pp. 1138–49, November 2003.

[16] S. Verdú, "Demodulation in the presence of multiuser interference: Progress and misconceptions," in D. Docampo, A. Figueiras, and F. Perez-Gonzalez, Eds., *Intelligent Methods in Signal Processing and Communications*, pp. 15–46, Birkhäuser, Boston, 1997.

[17] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[18] M. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication-Part I: System analysis," *IEEE Trans. Commun.*, pp. 795–9, August 1977.

[19] R. Pickholtz, L. Milstein, and D. Schilling, "Spread spectrum for mobile communications," *IEEE Trans. Veh. Tech.*, pp. 313–22, May 1991.

[20] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech.*, pp. 303–12, May 1991.

[21] H. Xiang, "Binary code-division multiple-access systems operating in multipath fading, noisy channels," *IEEE Trans. Commun.*, pp. 775–84, August 1985.

[22] S. Verdú, "Optimum multiuser signal detection," Ph.D. thesis, University of Illinois, Urbana Champaign, August 1984.

[23] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, pp. 85–96, January 1986.

[24] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Pers. Commun. Mag.*, pp. 46–58, April 1995.

[25] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multipleaccess channels," *IEEE Trans. Inform. Theory*, pp. 123–36, January 1989.

[26] M. L. Honig and H. V. Poor, "Adaptive interference suppression," in *Wireless Communications*: *Signal Processing Perspectives*, ch. 2, Prentice-Hall, Englewood Cliffs, NJ, 1998.

[27] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, pp. 622–40, March 1999.

[28] M. K. Varanasi and B. Aazhang, " Near-optimum detection in synchronous code division multipleaccess communications," *IEEE Trans. Commun.*, pp. 725–36, May 1991.

[29] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code division multipleaccess communications," *IEEE Trans. Commun.*, pp. 509–19, April 1990.

[30] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous CDMA," *IEEE Trans. Commun.*, pp. 285–90, February 1993.

[31] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous codedivision multipleaccess channels," *IEEE Trans. Commun.*, pp. 421–34, Febmary-April 1995.

[32] J. G. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Commun. Mag.*, pp. 19–29, April 2005.

[33] D. Divsalar, M. K. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Commun.*, pp. 258–68, February 1998.

[34] Y. C. Yoon, R. Kohno, and H. Imai, "A spread-spectrum multiaccess system with cochannel interference cancellation for multipath fading channels," *IEEE J. Sel. Areas Commun*., pp. 1067–75, September 1993.

[35] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Sel. Areas Commun*., pp. 796–807, June 1994.

[36] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread spectrum multiple-access channels," *IEEE J. Sel. Areas Commun*., pp. 641–9, May 1990.

[37] J. G. Andrews and T. H. Meng, "Optimum power control for successive interference cancellation with imperfect channel estimation," *IEEE Trans. Wireless Commun*., pp. 375–83, March 2003.

[38] X. Wang and H. V. Poor, " Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun*., pp. 91–103, January 1998.

[39] U. Mitra, "Comparison of maximum-likelihood-based detection for two multirate access schemes for CDMA signals," *IEEE Trans. Commun*., pp. 64–77, January 1999.

[40] X. Wang and H. V. Poor, *Wireless Communication Systems*: *Advanced Techniques for Signal Reception*, Prentice-Hall, Englewood Cliffs, NJ, 2004.

[41] R. Kohno, R. Meidan, and L. B. Milstein, "Spread spectrum access methods for wireless communications," *IEEE Commun. Mag*., pp. 58–67, January 1995.

[42] H. El Gamal and E. Geraniotis, "Comparing the capacities of FH/SSMA and DS/CDMA networks," *Proc. Internat. Sympos. Pers., Indoor, Mobile Radio Commun*., pp. 769–73, September 1998.

Chapter 14

Multiuser Systems

In multiuser systems the system resources must be divided among multiple users. This chapter develops techniques to allocate resources among multiple users and examines the fundamental capacity limits of multiuser systems. We know from Section 5.1.2 that signals of bandwidth B and time duration T occupy a signal space of dimension $2BT$. In order to support multiple users, the signal space dimensions of a multiuser system must be allocated to the different users.[1] Allocation of signaling dimensions to specific users is called *multiple access*.[2] Multiple access methods perform differently in different multiuser channels, and we will apply these methods to the two basic multiuser channels: downlink channels and uplink channels. Because signaling dimensions can be allocated to different users in an infinite number of different ways, multiuser channel capacity is defined by a *rate region* rather than a single number. This region describes all user rates that can be simultaneously supported by the channel with arbitrarily small error probability. We will discuss multiuser channel capacity regions for both the uplink and the downlink. We also consider random access techniques, whereby signaling dimensions are allocated only to active users, as well as power control, which ensures that users maintain the SINR required for acceptable performance. The performance benefits of multiuser diversity, which exploits the time-varying nature of the users' channels, is also described. We conclude with a discussion of the performance gains and signaling techniques associated with multiple antennas in multiuser systems.

14.1 Multiuser Channels: The Uplink and Downlink

A "multiuser" channel is any channel that must be shared among multiple users. There are two different types of multiuser channels, the *uplink* channel and the *downlink* channel, which are illustrated in Figure 14.1. A downlink, also called a broadcast channel or forward channel, has one transmitter sending to many receivers. Since the signals transmitted to all users originate from the downlink transmitter, the transmitted signal $s(t) = \sum_{k=1}^{K} s_k(t)$, with total power P and bandwidth B, is the sum of signals transmitted to all K users. Thus, the total signaling dimensions and power of the transmitted signal must be divided among the different users. Synchronization of the different users is relatively easy in the downlink because all signals originate from the same transmitter, although multipath in the channel can corrupt this synchronization. Another important characteristic of the downlink is that both signal and interference are distorted by the same channel. In particular, user k's signal $s_k(t)$ and all interfering signals $s_j(t)$, $j \neq k$, pass through user k's channel $h_k(t)$ to arrive at k's receiver. This is a fundamental difference between the uplink and the downlink, since in the uplink signals from different users are distorted by different

[1] Allocation of signaling dimensions through either multiple access or random access is performed by the "medium access control" layer in the open systems interconnect (OSI) network model [1, Chap. 1.3].

[2] The dimensions allocated to the different users need not be orthogonal, as in multiuser spread spectrum with nonorthogonal spreading codes or the superposition coding technique discussed in Section 14.6.

channels. Examples of wireless downlinks include all radio and television broadcasting, the transmission link from a satellite to multiple ground stations, and the transmission link from a base station to the mobile terminals in a cellular system.

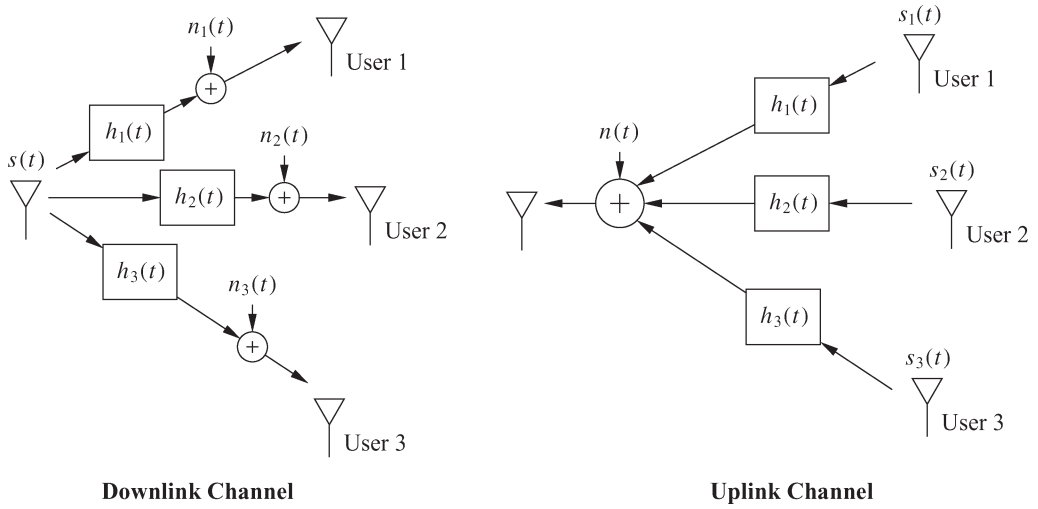


Figure 14.1: Downlink and uplink channels.

An uplink channel, also called a multiple access channel[3] or reverse channel, has many transmitters sending signals to one receiver, where each signal must be within the total system bandwidth B. However, in contrast to the downlink, in the uplink each user has an individual power constraint P_k associated with its transmitted signal $s_k(t)$. In addition, since the signals are sent from different transmitters, these transmitters must coordinate if signal synchronization is required. Figure 14.1 also indicates that the signals of the different users in the uplink travel through different channels, so even if the transmitted powers P_k are the same, the received powers associated with the different users will be different if their channel gains are different. Examples of wireless uplinks include laptop wireless LAN cards transmitting to a wireless LAN access point, transmissions from ground stations to a satellite, and transmissions from mobile terminals to a base station in cellular systems.

Most communication systems, including multiuser and point-to-point systems, are bi-directional with radios that can both transmit and receive communication signals. Bi-directional radios are referred to as transceivers. The radio transceiver that sends to users over a downlink channel and receives from these users over an uplink channel is often referred to as an access point or base station. Bi-directional multiuser systems typically separate the uplink and downlink channels into orthogonal signaling dimensions, typically using time or frequency dimensions. This separation is called *duplexing* or *half-duplexing*. Time-division duplexing (TDD) assigns orthogonal timeslots to a given user for receiving from an access point and transmitting to the access point, and frequency-division duplexing (FDD) assigns separate frequency bands for transmitting to and receiving from the access point. An advantage of TDD over FDD is that bi-directional channels are typically symmetrical in their channel gains, so channel measurements made in one direction can be used to estimate the channel in the other direction. This is not necessarily the case for FDD in frequency-selective fading: if the frequencies assigned to each direction are separated by more than the coherence bandwidth associated with the channel multipath, then these channels will exhibit independent fading. An alternative to half-duplex transmission is *full-duplex* transmission, whereby radios simultaneously transmit and receive data within the same frequency band. A full-duplex radio has double the spectral efficiency of a half-duplex radio and latency is also reduced by half relative to TDD. However, full-duplex radios suffer from significant *self-interference* since the transmitted signal emanating from the radio is typically at

[3]Note that multiple access techniques must be applied to both multiple access channels – that is, to uplinks as well as to downlinks.

much higher power than the received signal which has undergone path loss, shadowing, and fading. Full-duplex systems are discussed in more detail in Section 14.2.7.

14.2 Multiple Access

Efficient allocation of signaling dimensions between users is a key design aspect of both uplink and downlink channels, since bandwidth is usually scarce and/or very expensive. When dedicated channels are allocated to users it is often called multiple access.[4] Applications with continuous transmission and delay constraints, such as voice or video, typically require dedicated channels for good performance to ensure their transmission is not interrupted. Dedicated channels are obtained from the system signal space using a channelization method such as time division, frequency division, code division, or some combination of these techniques. Allocation of signaling dimensions for users with bursty transmissions generally use some form of random channel allocation that does not guarantee channel access. Bandwidth sharing using random channel allocation is called random multiple access or simply *random access*, which will be described in Section 14.3. In general, the choice of whether to use multiple access or random access – and which specific multiple or random access technique to apply – will depend on the system applications, the traffic characteristics of the users in the system, the performance requirements, and the characteristics of the channel and other interfering systems operating in the same bandwidth.

Multiple access techniques divide up the total signaling dimensions into channels and then assign these channels to different users. The most common methods to divide up the signal space are along the time, frequency, or code axes. The different user channels are then created by an orthogonal or non-orthogonal division along these axes: time-division multiple access (TDMA) and frequency-division multiple access (FDMA) are orthogonal channelization methods, whereas code-division multiple access (CDMA) can be either orthogonal or non-orthogonal, depending on the code design. Directional antennas, often obtained through antenna array processing, add an additional angular dimension that can also be used to channelize the signal space; this technique is called space-division multiple access (SDMA). Non-orthogonal multiple-access (NOMA) techniques include CDMA as well as other methods to create overlapping channels in time or frequency. Orthogonal channels can also be created in the delay-Doppler domain by assigning different delay-Doppler blocks to each user.

The performance of different multiple access methods depends on whether they are applied to an uplink or downlink as well as on their specific characteristics. The TDMA, FDMA, and orthogonal CDMA techniques are all equivalent in the sense that they orthogonally divide up the time-frequency signaling dimensions and hence create the same number of orthogonal channels. In particular, given a signal space of dimension $2BT$, it follows that N orthogonal channels of dimension $2BT/N$ can be created regardless of the channelization method. As a result, all multiple access techniques that divide the time-frequency signal space orthogonally have the same capacity when applied to additive white Gaussian noise (AWGN) channels, as will be discussed in Sections 14.6 and 14.7. However, channel impairments such as flat and frequency-selective fading affect these techniques in different ways, which leads to different channel capacities and different performance in typical wireless channels. There has been much work to analyze the relative performance of different multiple access techniques for current and future wireless systems (see e.g. [9, 12, 13, 14, 15, 16, 17]). Although analysis and general conclusions can be made for simple system and channel models, it is difficult to come up with a definitive answer as to the best technique for a complex multiuser system under a range of typical operating conditions. As with most engineering design questions, the choice of which multiple access technique to use will depend on the system requirements and characteristics along with cost and complexity constraints.

[4] An uplink channel is also referred to as a multiple access channel, but multiple access techniques are needed for both uplinks and downlinks.

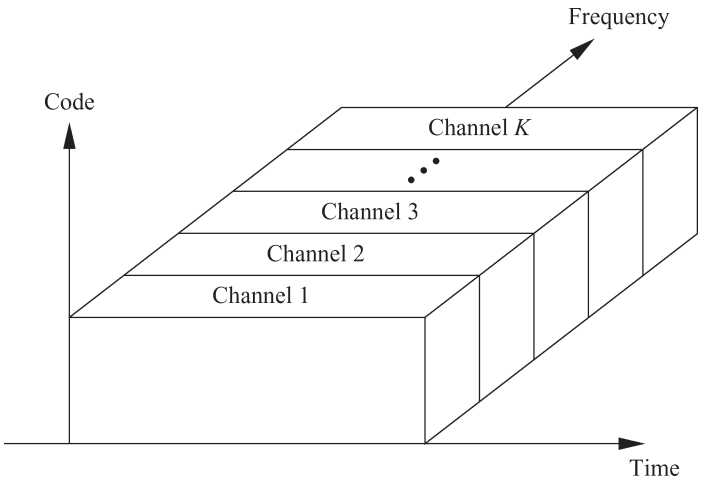


Figure 14.2: Frequency-division multiple access.

14.2.1 Frequency-Division Multiple Access (FDMA)

In FDMA the system signaling dimensions are divided along the frequency axis into nonoverlapping channels, and each user is assigned a different frequency channel; see Figure 14.2. The channels often have guard bands between them to compensate for imperfect filters, adjacent channel interference, and spectral spreading due to Doppler. If the channels are sufficiently narrowband, then the individual channels will not experience frequency-selective fading even if the total system bandwidth is large. Transmission is continuous over time, which can complicate overhead functions such as channel estimation because these functions must be performed simultaneously and in the same bandwidth as data transmission; FDMA also requires frequency-agile radios that can tune to the different carriers associated with the different channels. It is difficult to assign multiple channels to the same user under FDMA, since this requires the radios to simultaneously demodulate signals received over multiple frequency channels. Even so, FDMA is the most common multiple access option for analog communication systems, where transmission is continuous, and serves as the basis for the AMPS and TACS analog cellular phone standards [2, Chap. 11.1]. Multiple access in OFDM systems, called OFDMA, implements FDMA by assigning different subcarriers to different users.

Example 14.1: First-generation analog systems were allocated a total bandwidth of $B = 25$ MHz for uplink channels and another $B = 25$ MHz for downlink channels. This bandwidth allocation was split between two operators in every region, so each operator had 12.5 MHz for both their uplink and downlink channels. Each user was assigned $B_c = 30$ kHz of spectrum for its analog voice signal, corresponding to 24 kHz for the FM modulated signal with 3-kHz guard bands on each side. The total uplink and downlink bandwidths also required guard bands of $B_g = 10$ kHz on each side in order to mitigate interference to and from adjacent systems. Find the total number of analog voice users that could be supported in the total 25 MHz of bandwidth allocated to the uplink and the downlink. Also consider a more efficient digital system, with high-level modulation (so that only 10-kHz channels are required for a digital voice signal) and with tighter filtering (so that only 5-kHz guard bands are required on the band edges). How many users can be supported in the same 25 MHz of spectrum for this more efficient digital system?

Solution: For either the uplink or the downlink, guard bands on each side of the users' channels means that a total bandwidth of $NB_c + 2B_g$ is required for N users. Thus, the total number of users that can be supported in the total uplink or downlink bandwidth $B = 25$ kHz is

$$N = \frac{B - 2B_g}{B_c} = \frac{25 \cdot 10^6 - 2 \cdot 10 \cdot 10^3}{30 \cdot 10^3} = 832, \tag{14.1}$$

or 416 users per operator. Indeed, first-generation analog systems could support 832 users in each cell. The digital system has

$$N = \frac{B - 2B_g}{B_c} = \frac{25 \cdot 10^6 - 2 \cdot 5 \cdot 10^3}{10 \cdot 10^3} = 2499 \tag{14.2}$$

users that can be supported in each cell, a threefold increase over the analog system. The increase is primarily due to the bandwidth savings of the high-level digital modulation, which can accommodate a voice signal in one third the bandwidth of the analog voice signal.

14.2.2 Time-Division Multiple Access (TDMA)

In TDMA, the system dimensions are divided along the time axis into nonoverlapping channels, and each user is assigned a different cyclically repeating timeslot; see Figure 14.3. These TDMA channels occupy the entire system bandwidth, which is typically wideband, so some form of ISI mitigation is required. The cyclically repeating timeslots imply that transmission is not continuous for any user. Therefore, digital transmission techniques that allow for buffering are required. The fact that transmission is not continuous simplifies overhead functions such as channel estimation, since these functions can be performed during the timeslots occupied by other users. Time-division multiple access also has the advantage that it is simple to assign multiple channels to a single user by simply assigning him multiple timeslots.

A major difficulty of TDMA, at least for uplink channels, is the requirement for synchronization among the different users. Specifically, in a downlink channel all signals originate from the same transmitter and pass through the same channel to any given receiver. Thus, for flat-fading channels, if users transmit on orthogonal timeslots then the received signal will maintain this orthogonality. However, in the uplink channel the users transmit over different channels with different respective delays. To maintain orthogonal timeslots in the received signals, the different uplink transmitters must synchronize such that, *after* transmission through their respective channels, the received signals are orthogonal in time. This synchronization is typically coordinated by the base station or access point, and it can entail significant overhead. Multipath can also destroy time-division orthogonality in both uplinks and downlinks if the multipath delays are a significant fraction of a timeslot. Hence TDMA channels often have guard bands between them to compensate for synchronization errors and multipath. Another difficulty of TDMA is that, with cyclically repeating timeslots, the channel characteristics change on each cycle. Thus, receiver functions that require channel estimates, like equalization, must re-estimate the channel on each cycle. When transmission is continuous the channel can be tracked, which is more efficient. Time-division multiple access is used in the GSM, PDC, and IS-136 digital cellular phone standards [2, Chap. 11].

Example 14.2: The original GSM design uses 25 MHz of bandwidth for the uplink and for the downlink, the same as AMPs. This bandwidth is divided into 125 TDMA channels of 200 kHz each. Each TDMA channel consists of eight user timeslots; the eight timeslots along with a preamble and trailing bits form a *frame*, which is cyclically repeated in time. Find the total number of users that can be supported in the GSM system and the channel bandwidth of each user. If the root mean square delay spread of the channel is 10 μs, will intersymbol interference mitigation be needed in this system?

Solution: Since there are eight users per channel and 125 channels, the total number of users that can be supported in this system is $125 \cdot 8 = 1000$ users. The bandwidth of each TDMA channel is $25 \cdot 10^6/125 = 200$ kHz. A delay

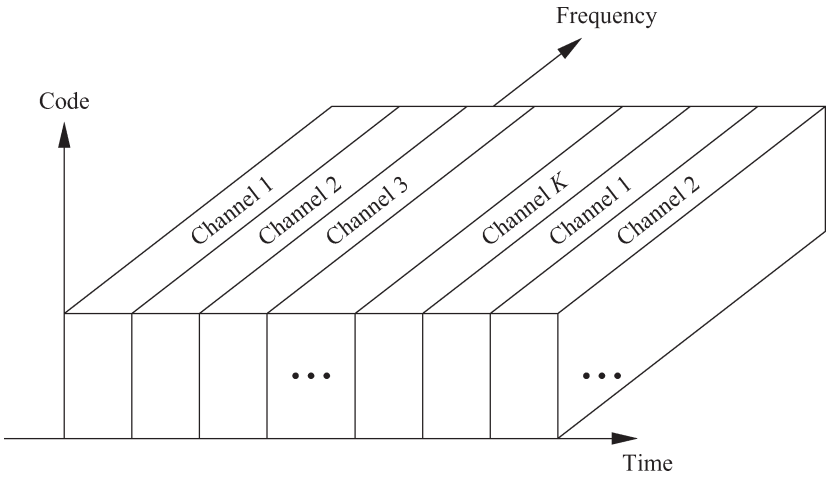


Figure 14.3: Time-division multiple access.

spread of 10 μs corresponds to a channel coherence bandwidth of $B_c \approx 100$ kHz, which is less than the TDMA channel bandwidth of 200 kHz. Thus, ISI mitigation is needed. The GSM specification includes an equalizer to compensate for ISI, but the type of equalizer is at the discretion of the designer.

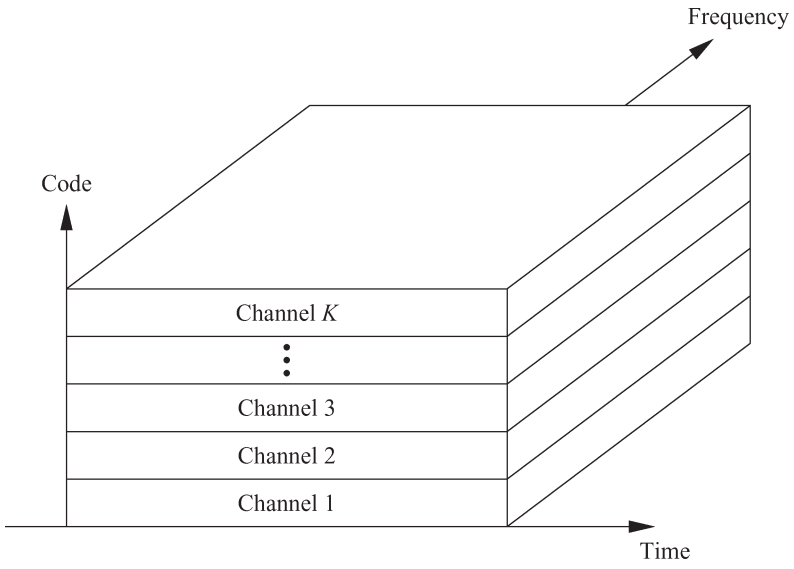


Figure 14.4: Code-division multiple access.

14.2.3 Code-Division Multiple Access (CDMA)

In CDMA the information signals of different users are modulated by orthogonal or nonorthogonal spreading codes. The resulting spread signals simultaneously occupy the same time and bandwidth, as shown in Figure 14.4. The receiver uses the spreading code structure to separate out the different users. The most common form of CDMA is multiuser spread spectrum with either direct sequence or frequency hopping, which were described and analyzed in Sections 13.4 and 13.5.

Downlinks typically use orthogonal spreading codes such as Walsh–Hadamard codes, although the orthogonality can be degraded by multipath. Uplinks generally use nonorthogonal codes owing to the difficulty of user synchronization and the complexity of maintaining code orthogonality in uplinks with multipath [6]. One of the big advantages of nonorthogonal CDMA in uplinks is that little dynamic coordination of users in time or frequency is required, since the users can be separated by the code properties alone. In addition, since TDMA and FDMA carve up the signaling dimensions orthogonally, there is a hard limit on how many orthogonal channels can be obtained. This is also true for CDMA using orthogonal codes, but if nonorthogonal codes are used then there is no hard limit on the number of channels that can be obtained. However, because nonorthogonal codes cause mutual interference between users, the more users that simultaneously share the system bandwidth using nonorthogonal codes, the higher the level of interference, which degrades system performance for all users. A nonorthogonal CDMA scheme also requires power control in the uplink to compensate for the near–far effect. The near–far effect arises in the uplink because the channel gain between a user's transmitter and the receiver is different for different users. Specifically, suppose that one user is very close to his base station or access point while another user is far away. If both users transmit at the same power level, then the interference from the close user will swamp the signal from the far user. Thus, power control is used such that the received signal power of all users is roughly the same. This form of power control, which essentially inverts any attenuation and or fading on the channel, causes each interferer to contribute an equal amount power, thereby eliminating the near–far effect. Code-division multiple access systems with nonorthogonal spreading codes can also use MUD to reduce interference between users. Multiuser detection provides considerable performance improvement even under perfect power control, and it works even better when the power control is jointly optimized with the MUD technique [7]. We will see in Sections 14.6 and 14.7 that a form of CDMA with multiuser detection achieves the Shannon capacity of both the uplink and the downlink, although the capacity-achieving transmission and reception strategies for the two channels are quite different. Finally, it is simple to allocate multiple channels to one user with CDMA by assigning that user multiple codes. Code-division multiple access is used for multiple access in the IS-95 digital cellular standards, with orthgonal spreading codes on the downlink and a combination of orthogonal and nonorthogonal codes on the uplink [2, Chap. 11.4]. It is also used in the W-CDMA and CDMA2000 digital cellular standards [8, Chap. 10.5].

Example 14.3: The signal-to-interference power ratio (SIR) for a CDMA uplink with nonorthogonal codes under the standard Gaussian assumption was given in (13.54) as

$$\mathrm{SIR} = \frac{3G}{K-1}, \tag{14.3}$$

where K is the number of users and $G \approx 128$ is the ratio of spread bandwidth to signal bandwidth. In IS-95 the uplink channel is assigned 1.25 MHz of spectrum. Thus, the bandwidth of the information signal prior to spreading is $B_s \approx 1.25 \cdot 10^6/128 = 9.765$ kHz. Neglecting noise, if the required SIR on a channel is 10 dB, how many users can the CDMA uplink support? How many could be supported within the same total bandwidth for an FDMA system?

Solution: To determine how many users can be supported, we invert the SIR expression to get

$$K \leq \frac{3G}{\mathrm{SIR}} + 1 = \frac{384}{10} + 1 = 39.4; \tag{14.4}$$

since K must be an integer, the system can support 39 users. In FDMA we have

$$K = \frac{1.25 \cdot 10^6}{9.765 \cdot 10^3} = 128, \tag{14.5}$$

so the total system bandwidth of 1.25 MHz can support 128 channels of 9.765 kHz. This calculation implies that FDMA is three times more efficient than nonorthogonal CDMA under the standard Gaussian assumption for code

cross-correlation (FDMA is even more efficient under different assumptions about the code cross-correlation). But, in fact, IS-95 typically supports 64 users on the uplink and downlink by allowing variable voice compression rates depending on interference and channel quality and by taking advantage of the fact that interference is not always present (called a voice-activity factor). Although this makes CDMA less efficient than FDMA for a single cell, cellular systems incorporate channel reuse, which can be done more efficiently in CDMA than in FDMA (see Section 15.1.4).

14.2.4 Space-Division Multiple Access (SDMA)

Space-division multiple access uses direction (angle) as another dimension in signal space, which can be channelized and assigned to different users, as shown in Figure 14.5. SDMA is typically implemented using a sectorized antenna or an antenna array. For the sectorized antenna, the 360° angular range is divided into K sectors. There is high directional gain in each sector and little interference between sectors. If there are more than K users, TDMA or FDMA can be used to channelize users within a sector. When a user moves between sectors, it must be handed off to the new sector. An alternative SDMA technique to sectorization is an antenna array. An antenna array can create static spatial dimensions or dynamically assign spatial channels to different users as their locations and channel conditions change by adjusting the antenna array parameters. An antenna array with K elements can create K spatially orthogonal channels. If each of the K users in the system has a single antenna, then each user gets one spatial channel. As with sectorized antennas, TDMA or FDMA can be used to create multiple orthogonal channels within one spatial channel. Multiuser systems where both the transmitter and receiver have antenna arrays are called *Multiuser MIMO Systems*. As described in more detail in Section 14.10, for a multiuser MIMO system with an antenna array of size K, a user with $N \leq K$ antennas can support up to N independent spatial streams within the kth spatial channel associated with the K-element antenna array. Multiuser MIMO systems must adjust the antenna parameters at both the transmitter and receiver arrays as channel conditions and user locations change to fully utilize the available spectral dimensions.

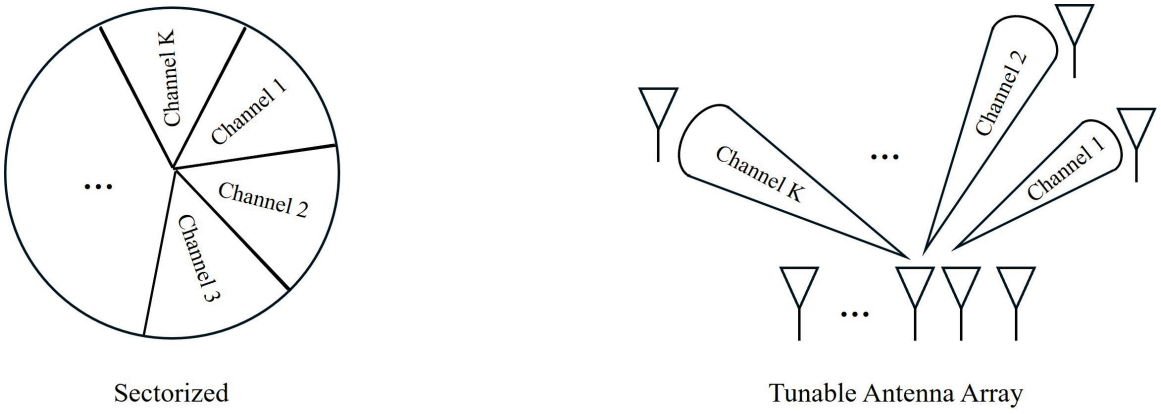


Figure 14.5: Space-Division Multiple Access

14.2.5 Non-orthogonal Multiple Access (NOMA)

As will be described in Sections 14.6-14.7, the channel capacity of a multiuser channel for both the downlink and the uplink are achieved through non-orthogonal multiple access (NOMA) techniques. In particular, superposition coding with multiuser decoding achieves the downlink channel capacity, while superimposing user signals within

the same channel and separating them at the receiver using multiuser detection achieves the uplink channel capacity. These fundamental limits indicate that NOMA techniques may achieve higher spectral efficiency than orthogonal multiple access methods. In addition, as seen in Section 14.2.3 for non-orthogonal CDMA, benefits of NOMA over orthogonal multiple access methods include limited user coordination requirements, no hard limit on the number of user channels, and the potential to remove most if not all of the interference between users with multiuser detection.

The primary techniques used for NOMA fall within two categories: non-orthogonal CDMA, described in Section 14.2.3, which is referred to as *code-domain NOMA*, and *power-domain NOMA*, which uses different received power levels of different users along with successive interference cancellation (SIC) to separate the signals of users sharing the same channel. Specifically, in power-domain NOMA a set of N users are assigned to the same channel. The received signal at an uplink or downlink receiver is the superposition of all user signals sharing this channel. The received power of the nth user at the receiver is determined by its transmit power and the channel path loss, shadowing, and multipath fading between its transmitter and the receiver. The received power of the N different users can thus be set to a specific set of values $(P_1, \ldots, P_N)$ by setting the corresponding transmit power of the nth user to the desired received power P_n of that user plus the power loss introduced by that user's channel. SIC at the receiver then works as follows; the highest-power user, i.e. the user for which $P_i > P - j, \forall j \neq i$, is decoded first based on treating all other user signals as noise. The signal of this first user is then subtracted out of the received signal and the second highest-power user is decoded by treating all lower-power users as noise. Then the signals of both the highest and second-highest power users are subtracted from the received signal in order to decode the third-highest-power user, and the process continues until all user signals have been decoded. In this process, the last signal decoded, corresponding to the lowest power user, has all interference removed before decoding, assuming the SIC can perfectly remove interference at each iteration stage.

The SINR of each user in power-domain NOMA with perfect SIC is dictated by the transmit powers of all users, their channel characteristics, and the order in which a given user is decoded, since only lower-powered users cause interference to that user. Due to these restrictions, not all SINRs $(\gamma_1, \ldots, \gamma_N)$ in an N-user power-domain NOMA system can be achieved. The set of SINRs that can be achieved, which is based on *power control* at each transmitter, will be analyzed in Section 14.4. Which set of SINRs utilized among those that are possible is a design parameter that incorporates performance tradeoffs between the users, including minimum rate requirements as well as fairness. These tradeoffs will also be analyzed in Section 14.4.

14.2.6 Hybrid Techniques

Hybrid techniques for multiple access use a combination of different orthogonal and non-orthogonal multiple access methods to either create additional channels or to improve performance over a single multiple access method. The combination of SDMA with TDMA or FDMA described in the previous section is an example of hybrid multiple access. NOMA techniques are often combined with either TDMA or FDMA to minimize the number of interfering users in a given channel. The most common hybrid multiple access technique combines orthogonal time and frequency division, which creates orthogonal time-frequency channels, often referred to as *time-frequency blocks*, within the allocated signal space. This hybrid technique is often implemented using time-division and OFDM, whereby the channels associated with different OFDM subcarriers are channelized in time to create time-frequency blocks which are then assigned to different users. This implementation, illustrated in Fig. 14.6, is referred to as Orthogonal Frequency Division Multiple Access (OFDMA). OFDMA can also be combined with tone hopping over different subcarriers to improve frequency diversity [9]. A similar idea can be applied to OTFS, whereby different delay-Doppler blocks can be assigned to different users [10, 11].

Hybrid methods are also applied to channels created through CDMA to either improve performance or create more flexibility for channel assignment. For example, direct-sequence spread spectrum (DSSS) can be combined with FDMA to break the system bandwidth into subbands; in this hybrid method, different users are assigned to different subbands with their signals spread across the subband bandwidth. Within a subband, the processing gain

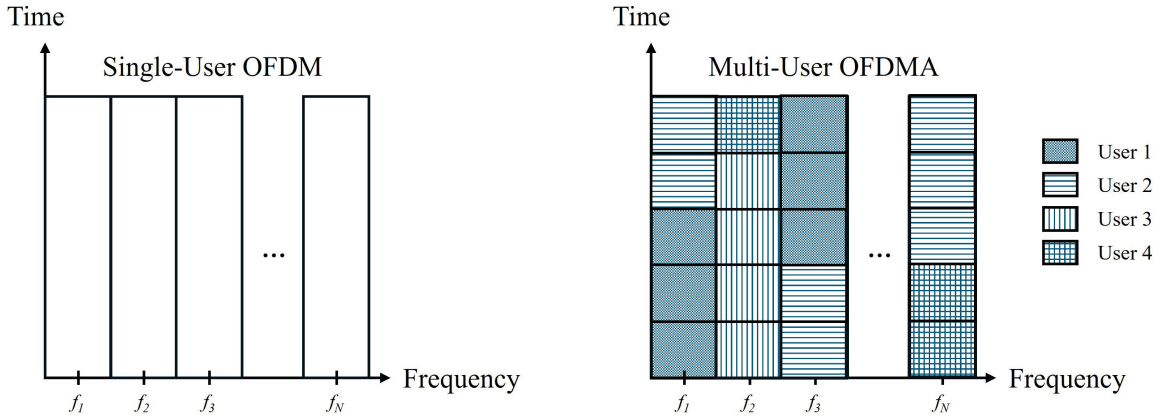


Figure 14.6: Multi-User OFDMA versus Single-User OFDM. In OFDMA, time-frequency blocks are created by applying time-division to OFDM subcarrier channels, which are then assigned to different users.

is smaller than it would be over the entire system bandwidth, so interference and ISI attenuation are reduced. In exchange for this performance loss, this hybrid technique does not require contiguous spectrum between subbands, and it also allows more flexibility in spreading user signals over different-size subbands depending on their requirements. Another hybrid method combines DS-CDMA with FH-CDMA so that the carrier frequency of the spread signal is hopped over the available bandwidth. This reduces the near–far effect because the interfering users change on each hop. Alternatively, TDMA and FH can be combined so that a channel with deep fading or interference is used only on periodic hops, enabling the mitigation of fading and interference effects via error correction coding. This idea was used in the GSM standard, which combined FH with its TDMA scheme in order to reduce the effect of strong interferers in other cells.

14.2.7 Full-Duplex Systems

In full-duplex systems radios transmit and receive simultaneously using the same bandwidth. A full-duplex radio requires one channel for transmission and reception rather than the two channels required by half-duplex systems, thereby doubling spectral efficiency. Latency is also reduced by half relative to TDD half-duplex systems. These full-duplex benefit come at the expense of the self-interference caused by the transmitted signal to the received signal, as illustrated in Fig. 14.7. The received signal, having experienced path loss, shadowing, and fading, is typically at much lower power than the transmitted signal generated by the power amplifier (PA). This analog signal can "leak" into any part of the analog receiver chain, causing transmitter self-interference. The resulting low SINR of the received signal makes reliable detection of this signal challenging. In fact, if the dynamic range of the A/D converter is smaller than the dynamic range between the self-interference and the received signal, then the received signal is completely eliminated in the receiver A/D conversion.

As shown in Fig. 14.8, analog signal processing of the self-interference prior to the A/D is typically required in full-duplex systems to reduce the self-interference to within the A/D dynamic range. Following analog interference reduction the signal is converted to the digital domain for further interference reduction. Since DSP techniques are generally more effective than analog techniques, most full-duplex systems utilize both analog and digital self-interference reduction methods: the analog methods which operate on the analog self-interference to get it to within the A/D dynamic range, and then digital methods to reduce the self-interference further in order to achieve the SINR required to reliably detect the received signal. Ideally the self-interference can be reduced below the noise floor of the receiver prior to baseband demodulation, in which case there is no effect of the self-interference

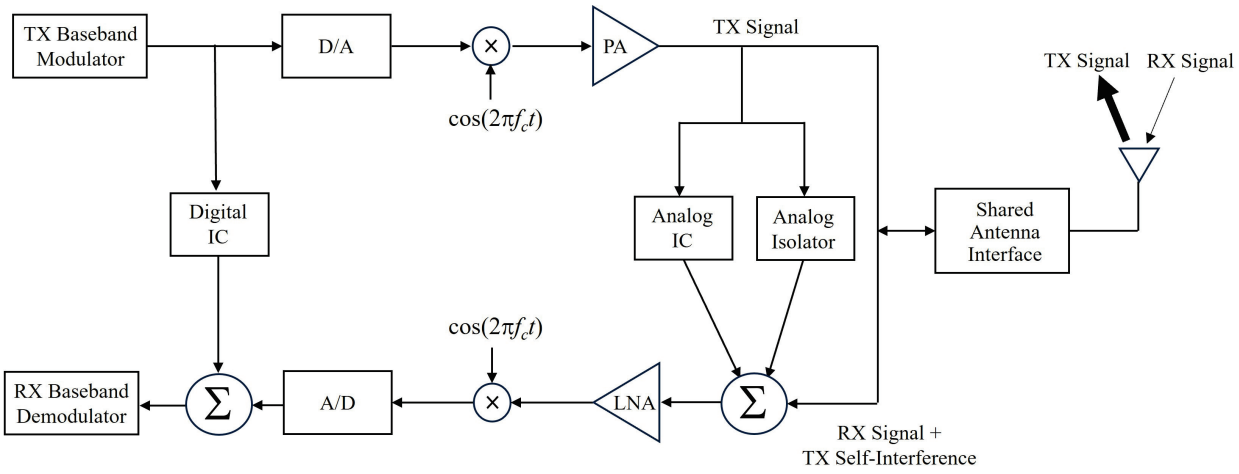


Figure 14.7: Full Duplex System

in detection of the received signal.

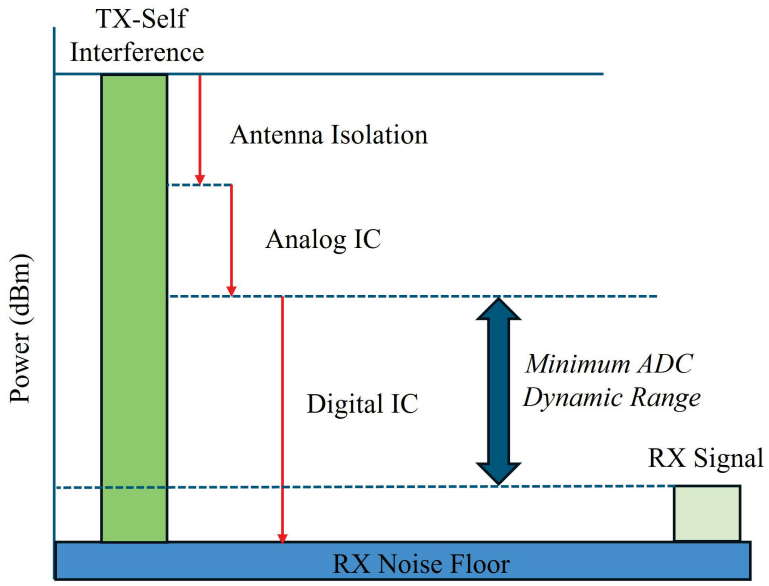


Figure 14.8: Analog and Digital Interference Reduction Stages

There are two main methods for analog self-interference reduction: isolation and interference cancellation (IC). Analog isolation uses a system of passive RF hardware components to isolate the receive signal from the transmit signal. A common design of these analog *isolators* is to introduce a tapped delay line with multiple phase shifts and corresponding attenuations applied to the analog transmit signal. Since the self-interference typically consists of multiple transmit signal components with different delays and attenuations, the isolator can partially or, ideally, fully cancel out the self-interference. However, due to imperfections in the analog hardware and unknown delays and attenuations of the components in the self-interference, isolators cannot typically remove enough interference to within the A/D dynamic range. In this case analog IC is needed. Analog IC reconstructs the analog transmit signal from the transmit baseband signal in order to cancel it out. This requires accurate modeling and

generation of the transmitter front end effects, including nonlinearities in the D/A converter and power amplifier (PA). Once the front-end channel used to generate the self-interference is accurately modeled, the self-interference can be accurately estimated by the receiver given that the transmitted signal sent through the front end channel is known. Successive interference cancellation, similar to that used in CDMA systems as described in Section 13.4.4, can be used in full-duplex as well, where each iteration removes some of the self-interference.

Multi-antenna full-duplex systems have one or more antennas dedicated to the transmitted signal and one or more dedicated to the received signal. Analog isolators separate the transmit antenna(s) from the receive antenna(s), and beamforming applied to transmit and receive arrays are used to separate the direction in which the transmitted signal is sent from the direction in which the received signal is received. The interference reduction that results depends on the isolator capabilities and the antenna array sizes.

Full-duplex radios were considered infeasible for commercial use prior to 2010 due to limitations in analog interference reduction and the overall system complexity. A wave of breakthroughs in that timeframe made full-duplex radios viable [3, 4], although they are not yet competitive with half-duplex radios in terms of cost and performance. More details on full-duplex systems and their self-interference mitigation techniques can be found in [5].

14.3 Random Access

Multiple access techniques are primarily for continuous applications like voice and video, where a dedicated channel facilitates good performance. However, most data applications do not require continuous transmission: data are generated at random time instances, so dedicated channel assignment can be extremely inefficient. Moreover, most systems have many more total users (active plus idle users) than can be accommodated simultaneously, so at any given time channels can only be allocated to users that need them. Random access strategies are used in such systems to efficiently assign channels to the active users.

All random access techniques are based on the premise of packetized data or *packet radio*. In packet radio, user data is collected into packets of N bits, which may include error detection/correction and control bits. Once a packet is formed it is transmitted over the channel. Assuming a fixed channel data rate of R bps, the transmission time of a packet is $\tau = N/R$. The transmission rate R is assumed to require the entire signal bandwidth, and all users transmit their packets over this bandwidth with no additional signaling to separate simultaneously transmitted packets. Thus, if packets from different users overlap in time a *collision* occurs, in which case both packets may be decoded unsuccessfully. Packets may also be decoded in error as a result of noise or other channel impairments. The probability of a packet decoding error is called the *packet error rate*. Analysis of random access techniques typically assumes that, collectively, the users accessing the channel generate packets according to a Poisson process at a rate of λ packets per unit time; that is, λ is the average number of packets that arrive in any time interval $[0,\ t]$ divided by t. Equivalently, λN is the average number of bits generated in any time interval $[0,\ t]$ divided by t. For a Poisson process, the probability that the number of packet arrivals in a time period $[0,\ t]$, denoted as $X(t)$, is equal to some integer k is given by

$$p(X(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \tag{14.6}$$

Poisson processes are memoryless, so that the number of packet arrivals during any given time period does not affect the distribution of packet arrivals in any other time period. Note that the Poisson model is not necessarily a good model for all types of user traffic – especially Internet data, where bursty data causes correlated packet arrivals [18].

The traffic *load* on the channel given Poisson packet arrivals at rate λ and packet transmission duration τ is defined as $L = \lambda\tau$. If the channel data rate is R_p packets per second then $\tau = 1/R_p = N/R$ for R the

channel data rate in bits per second. Note that L is unit-less: it is the ratio of the packet arrival rate divided by the packet rate that can be transmitted over the channel at the channel's data rate R. We assume that colliding packets are always decoded in error. Thus, if $L > 1$ then on average more packets (or bits) arrive in the system over a given time period than can be transmitted in that period, so systems with $L > 1$ are unstable. If the transmitter is informed by the receiver about packets received in error and retransmits these packets, then the packet arrival rate λ and corresponding load $L = \lambda\tau$ is computed based on arrivals of both new packets and packets that require retransmission. In this case L is referred to as the *total offered load*.

Performance of random access techniques is typically characterized by the *throughput* T of the system. The throughput, which is unitless, is defined as the ratio of the average rate of packets successfully transmitted divided by the channel packet rate R_p. The throughput thus equals the offered load multiplied by the probability of successful packet reception, $T = Lp$ (successful packet reception), where this probability is a function of the random access protocol in use as well as the channel characteristics, which can cause packet errors in the absence of collisions. Thus $T \leq L$. Also, since a system with $L > 1$ is unstable, stable systems have $T \leq L \leq 1$. Observe that the throughput is independent of the channel data rate R, since the load and corresponding throughput are normalized with respect to this rate. This allows analysis of random access protocols to be generic to any underlying link design or channel capacity. For a packet radio with a link data rate of R bps, the *effective data rate* of the system is RT, since T is the fraction of packets or bits successfully transmitted at rate R. The goal of a random access method is to make T as large as possible in order to fully utilize the underlying link rates. Note that in some circumstances overlapping packets do not cause a collision. In particular, short periods of overlap between colliding packets, different channel gains on the received packets, and/or error correction coding can allow one or more packets to be successfully received even with a collision. This is called the *capture effect* [19, Chap. 4.3].

Random access techniques were pioneered by Abramson with the ALOHA protocol [20], where data is packetized and users send packets whenever they have data to send. ALOHA is extremely inefficient owing to collisions between users, which leads to very low throughput. The throughput can be doubled by slotting time and synchronizing the users, but even then collisions lead to relatively low throughput values. Modifications of ALOHA protocols to avoid collisions and thereby increase throughput include carrier sensing, collision detection, and collision avoidance. Long bursts of packets can be scheduled to avoid collisions, but this typically takes additional overhead. In this section we will describe these various techniques for random access, their performance, and their design trade-offs.

14.3.1 Pure ALOHA

In pure or unslotted ALOHA, users transmit data packets as soon as they are formed. If we neglect the capture effect, then packets that overlap in time are assumed to be received in error and must be retransmitted. If we also assume packets that do not collide are successfully received (i.e., if there is no channel distortion or noise), then the throughput equals the offered load multiplied by the probability of no collisions: $T = Lp$ (no collisions). Suppose a given user transmits a packet of duration τ during time $[0,\ \tau]$. Then, if any other user generates a packet during time $[-\tau,\ \tau]$, that packet (of duration τ) will overlap with the transmitted packet, causing a collision. The probability of no collisions thus equals the probability that no packets are generated during the time $[-\tau,\ \tau]$. This probability is given by (14.6) with $t = 2\tau$:

$$p(X(t) = 0) = e^{-2\lambda\tau} = e^{-2L}, \tag{14.7}$$

with corresponding throughput

$$T = Le^{-2L}. \tag{14.8}$$

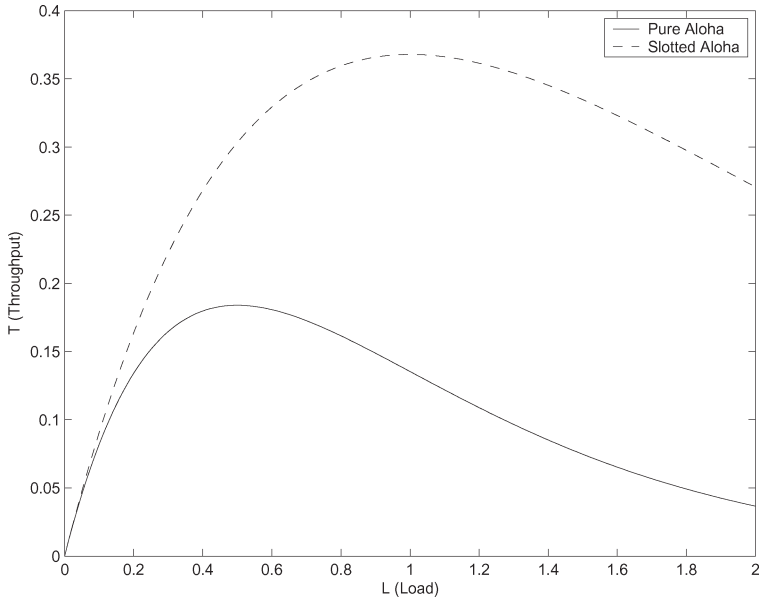


Figure 14.9: Throughput of pure and slotted ALOHA.

This throughput is plotted in Figure 14.9, where we see that throughput increases with offered load up to a maximum throughput of approximately .18 for $L = .5$, after which point it decreases. In other words, the data rate is only 18% of what it would be with a single user transmitting continuously on the system. The reason for this maximum is that, for small values of L, there are many idle periods when no user is transmitting, so throughput is small. As L increases, the channel is utilized more but collisions also start to occur. At $L = .5$ there is the optimal balance between users generating enough packets to utilize the channel with reasonable efficiency and these packet generations colliding infrequently. Beyond $L = .5$ the collisions become more frequent, which degrades throughput below its maximum; as L grows very large, most packets experience collisions and throughput approaches zero.

Part of the reason for the inefficiency of pure ALOHA is the fact that users can start their packet transmissions at any time, and any partial overlap of two or more packets destroys the successful reception of all packets. By synchronizing users such that all packet transmissions are aligned in time, the partial overlap of packet transmissions can be avoided. That is the basic premise behind slotted ALOHA.

14.3.2 Slotted ALOHA

In slotted ALOHA, time is assumed to be slotted in timeslots of duration τ, and users can only start their packet transmissions at the beginning of the next timeslot after the packet has formed. Thus, there is no partial overlap of transmitted packets, which increases throughput. Specifically, a packet transmitted over the time period $[0, \tau]$ is successfully received if no other packets are transmitted during this period. This probability is obtained from (14.6) with $t = \tau : p(X(t) = 0) = e^{-L}$, with corresponding throughput

$$T = Le^{-L}. \tag{14.9}$$

This throughput is also plotted in Figure 14.9, where we see that throughput increases with offered load up to a maximum of approximately $T = .37$ for $L = 1$, after which point it decreases. Thus, slotted ALOHA has double the maximum throughput as pure ALOHA, and it achieves this maximum at a higher offered load. Although this represents a marked improvement over pure ALOHA, the effective data rate is still less than 40% of the raw

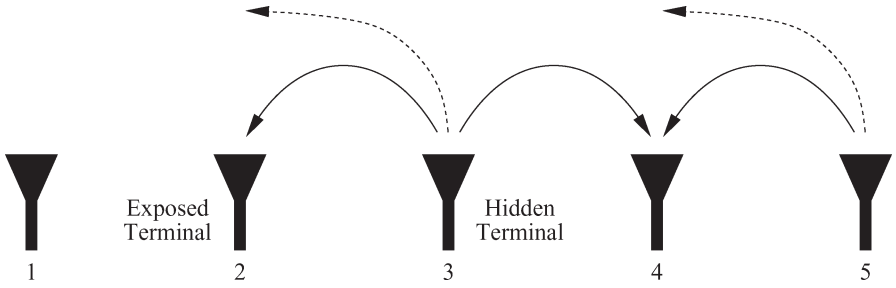


Figure 14.10: Hidden and exposed terminals.

transmission rate. This is extremely wasteful of the limited wireless bandwidth, so more sophisticated techniques are needed to increase efficiency.

Note that slotted ALOHA requires synchronization of all nodes in the network, which can entail significant overhead. Even in a slotted system, collisions occur whenever two or more users attempt transmission in the same slot. Error control coding can result in correct detection of a packet even after a collision, but if the error correction is insufficient then the packet must be retransmitted. A study on design optimization between error correction and retransmission is described in [21].

Example 14.4: Consider a slotted ALOHA system with a transmission rate of $R = 10$ Mbps. Suppose packets consist of 1000 bits. For what packet arrival rate λ will the system achieve maximum throughput, and what is the effective data rate associated with this throughput?
Solution: The throughput T is maximized for $L = \lambda\tau = 1$, where λ is the packet arrival rate and τ is the packet duration. With a 10-Mbps transmission rate and 1000 bits per packet, $\tau = 1000/10^6 = .1$ ms. Thus, $\lambda = 1/.0001 = 10^4$ packets per second maximizes throughput. The throughput for $L = 1$ is $T = .37$, so the effective data rate is $TR = 3.7$ Mbps. Thus, the data rate is reduced by roughly a factor of 3 (as compared to continuous data transmission) owing to the random nature of the packet arrivals and their corresponding collisions.

14.3.3 Carrier-Sense Multiple Access (CSMA)

Collisions can be reduced by carrier-sense multiple access, where users sense the channel and delay transmission if they detect that another user is currently transmitting. To be effective, detection time and propagation delays in the system must be small [22, Chap. 4.19]. After sensing a busy channel, a user typically waits a random time period before transmitting. This *random backoff* precludes multiple users simultaneously transmitting as soon as the channel is free. Carrier-sense multiple access works only when all users can detect each other's transmissions and the propagation delays are small. Wired LANs have these characteristics, so CSMA is part of the Ethernet protocol. However, the nature of the wireless channel may prevent a given user from detecting the signals transmitted by all other users. This gives rise to the *hidden terminal problem* (illustrated in Figure 14.10), whereby each node can hear its immediate neighbor but no other nodes in the network. In the figure, node 3 and node 5 each wish to transmit to node 4. Suppose node 5 starts its transmission. Since node 3 is too far away to detect this transmission, it assumes that the channel is idle and begins its transmission, thereby causing a collision with node 5's transmission. Node 3 is said to be "hidden" from node 5 because it cannot detect node 5's transmission.

ALOHA with CSMA also creates inefficiencies in channel utilization from the *exposed terminal problem*, also illustrated in Figure 14.10. Suppose the exposed terminal in this figure – node 2 – wishes to send a packet to node 1 at the same time that node 3 is sending to node 4. When node 2 senses the channel it will detect node 3's

transmission and assume the channel is busy, even though node 3 does not interfere with the reception of node 2's transmission by node 1. Thus node 2 will not transmit to node 1 even though no collision would have occurred. Exposed terminals occur only in multihop networks, so we will defer their discussion until Chapter 16.

The collisions introduced by hidden terminals are often avoided in wireless networks by a four-way handshake prior to transmission [23, 24]. This *collision avoidance* is accomplished as follows. A node that wants to send a data packet will first wait for the channel to become available and then transmit a short RTS (request to send) packet. The potential receiver, assuming it perceives an available channel, will immediately respond with a CTS (clear to send) packet that authorizes the initiating node to transmit and also informs neighboring hidden nodes (i.e., nodes that are outside the communication range of the transmitter but within the communication range of the receiver) that they must remain silent for the duration of the transmission. Nodes that overhear the RTS or CTS packet will refrain from transmitting over the expected packet duration. A node can send an RTS packet only if it perceives an idle channel and has not been silenced by another control packet. A node will transmit a CTS packet only if it has not been silenced by another control packet. The RTS/CTS handshake is typically coupled with random backoff to avoid all nodes transmitting as soon as the channel becomes available. In some incarnations [24, 25], including the 802.11 WLAN standard [8, Chap. 14.3], the receiver sends an ACK (acknowledgement) packet back to the transmitter to verify when it has correctly received the packet, after which the channel again becomes available.

Another technique to avoid hidden terminals is busy-tone transmission. In this strategy, users first check to see whether the transmit channel is busy by listening for a "busy tone" on a separate control channel [1, Chap. 4.6]. There is typically not an actual busy tone; instead, a bit is set in a predetermined field on the control channel. This scheme works well in preventing collisions when a centralized controller can be "heard" by users throughout the network. In a flat network without centralized control, more complicated measures are used to ensure that any potential interferer on the first channel can hear the busy tone on the second [26, 27]. Hybrid techniques using handshakes, busy-tone transmission, and power control can also be used [27]. Collisions can also be reduced by combining DSSS with ALOHA. In this scheme, each user modulates his signal with the same spreading code, but if user transmissions are separated by more than a chip time, the interference due to a collision is reduced by the code autocorrelation [28].

14.4 Scheduling

Random access protocols work well with bursty traffic, where there are many more users than available channels yet these users rarely transmit. If users have long strings of packets or continuous-stream data, then random access works poorly because most transmissions result in collisions. In this scenario performance can be improved by assigning channels to users in a more systematic fashion through transmission scheduling. In scheduled access, the available bandwidth is channelized into multiple orthogonal or non-orthogonal channels using the techniques described in Section 14.2. Each node schedules its transmission on different channels in such a way as to avoid conflicts with neighboring nodes while making the most efficient use of the available signaling dimensions. The most common forms of scheduling in wireless systems are round robin and opportunistic scheduling. Round robin does not take into account channel characteristics whereas opportunistic scheduling utilizes user channel state information (CSI) in its scheduling mechanism. Scheduling is easier to implement on the downlink rather than the uplink since transmissions to all users are centrally controlled and synchronized in time. Scheduling in the uplink requires coordination across users, particularly for time synchronization. We will describe scheduling protocols based on downlink transmissions for simplicity; the uplink versions are similar but require an additional user coordination protocol. One such protocol is packet-reservation multiple access (PRMA).

In round robin scheduling users takes turns using a given channel for a fixed period of time T. Hence, if

there are N users, the first user transmits over time interval $[0, T]$, the second user over time interval $[T, 2T]$, and so on until all users have transmitted over the channel for one time interval. Then the process repeats. If there is more than one channel available to these users, then round robin can be used for each of the available channels. Round robin is a fair algorithm in that every user is assigned the same proportion of the available system resources. However, that does not mean each user is assigned the same data rate, since that rate will depend on the channel quality between the transmitter and that user's receiver. Round robin also does not take into account any prioritization of users.

An alternative to round robin scheduling maximizes total throughput across all users. This goal is achieved by assigning the channel to the user whose channel capacity during each time interval is the highest. This is a form of opportunistic scheduling called *MaxRate* scheduling. MaxRate scheduling requires CSI for of all users during each time interval to determine which one has the highest channel capacity. The problem with MaxRate scheduling is that users with poor channels are never able to transmit. This problem is addressed through another form of opportunistic scheduling called *proportional fair scheduling*. In proportional fair scheduling, over a given time interval T, the channel is assigned to the user whose ratio of channel capacity C_i to past throughput T_i is maximized. This allows users with poor channels to transmit whenever their past throughput drops significantly since, as T_i shrinks due to other users accessing the channel over previous intervals, the ratio C_i/T_i grows and will eventually exceed that of all other users. MaxRate and proportionally fair scheduling are two simple examples of opportunistic scheduling, which in general assigns the channel over each time interval T to the user with the highest value of the metric associated with the scheduler design. The metric can take into account multiple properties of the system such as minimum or maximum user throughput, total throughput, fairness, delay, and jitter. When multiple channels are available, opportunistic scheduling is based on maximizing a two-dimensional metric over both time and frequency. A detailed discussion of opportunistic scheduling and the different metrics it uses can be found in [29].

For uplink channels, even with a scheduling access protocol, some form of ALOHA will still be needed because a predefined mechanism for scheduling is, by definition of random access, unavailable at startup. ALOHA provides a means for initial contact and the establishment of some form of scheduled access for the transmission of relatively large amounts of data. A systematic approach to this initialization that also combines the benefits of random access for bursty data with scheduling for continuous data is *packet-reservation multiple access* (PRMA) [30]. This technique assumes a slotted system with both continuous and bursty users (e.g., voice and data users). Multiple users vie for a given time slot under a random access strategy. A successful transmission by one user in a given timeslot reserves that timeslot for all subsequent transmissions by the same user. If the user has a continuous or long transmission then, after successfully capturing the channel, he has a dedicated channel for the remainder of his transmission (assuming subsequent transmissions are not corrupted by the channel; such corruption causes users to lose their slots and they must then recontend for an unreserved slot, which can entail significant delay and packet dropping [31]). When this user has no more packets to transmit, the slot is returned to the pool of available slots that users attempt to capture via random access. Thus, data users with short transmissions benefit from the random access protocol assigned to unused slots, and users with continuous transmissions get scheduled periodic transmissions after successfully capturing an initial slot. A similar technique using a combined reservation and ALOHA policy is described in [32].

14.5 Power Control

Power control is applied to systems where users interfere with each other. The goal of power control is to adjust the transmit powers of all users such that the SINR of each user meets a given threshold required for acceptable performance. This threshold may be different for different users, depending on their required performance. This problem is straightforward for the downlink, where users and interferers have the same channel gains, but is more

complicated in the uplink, where the channel gains may be different. Seminal work on power control for cellular systems and ad hoc networks was done in [33, 34, 35], and power control for the uplink is a special case for which these results can be applied. In the uplink model, the kth transmitter has a fixed channel power gain g_k to the receiver. The quality of each link is determined by the SINR at the intended receiver. In an uplink with K interfering users, we denote the SINR for the kth user as

$$\gamma_k = \frac{g_k P_k}{n + \rho \sum_{j \neq k} g_j P_j}, \quad k = 1, \ldots, K, \tag{14.10}$$

where P_k is the power of the kth transmitter, n is the receiver noise power, and ρ is interference reduction due to signal processing. For example, in a CDMA uplink the interference power is reduced by the processing gain of the code, so $\rho \approx 1/G$ for G the processing gain, whereas in TDMA $\rho = 1$.

Each link is assumed to have a minimum SINR requirement $\gamma_k^* > 0$. This constraint can be represented in matrix form with componentwise inequalities as

$$(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u} \quad \text{with } \mathbf{P} > 0, \tag{14.11}$$

where $\mathbf{P} = (P_1, P_2, \ldots, P_K)^T$ is the vector of transmitter powers,

$$\mathbf{u} = \left(\frac{n\gamma_1^*}{g_1}, \frac{n\gamma_2^*}{g_2}, \ldots, \frac{n\gamma_K^*}{g_K} \right)^T \tag{14.12}$$

is the vector of noise power scaled by the SINR constraints and channel gains, and $\mathbf{F}$ is a matrix with

$$F_{kj} = \begin{cases} 0 & k = j, \\ \gamma_k^* g_j \rho / g_k & k \neq j, \end{cases} \tag{14.13}$$

for $k, j = 1, 2, \ldots, K$.

The matrix $\mathbf{F}$ has nonnegative elements and is irreducible. Let ρ_F be the Perron–Frobenius eigenvalue of $\mathbf{F}$. This is the maximum modulus eigenvalue of $\mathbf{F}$, and for $\mathbf{F}$ irreducible this eigenvalue is simple, real, and positive. Moreover, from the Perron–Frobenius theorem and standard matrix theory [36], the following statements are equivalent.

1. $\rho_F < 1$.
2. There exists a vector $\mathbf{P} > 0$ (i.e., $P_k > 0$ for all k) such that $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$.
3. $(\mathbf{I} - \mathbf{F})^{-1}$ exists and is positive componentwise.

Furthermore, if any of these conditions holds then we also have that $\mathbf{P}^* = (\mathbf{I} - \mathbf{F})^{-1}\mathbf{u}$ is the Pareto optimal solution to (14.11). That is: if $\mathbf{P}$ is any other solution to (14.11), then $\mathbf{P} \geq \mathbf{P}^*$ componentwise. Hence, if the SINR requirements for all users can be met simultaneously then the best power allocation is $\mathbf{P}^*$, which will minimize the transmit power of the users.

In [34] the authors also show that the following iterative power control algorithm converges to $\mathbf{P}^*$ when $\rho_F < 1$ and diverges to infinity otherwise. This iterative Foschini–Miljanic algorithm is given by

$$\mathbf{P}(i+1) = \mathbf{F}\mathbf{P}(i) + \mathbf{u} \tag{14.14}$$

for $i = 1, 2, 3, \ldots$. Furthermore, this algorithm can be simplified to a per-user version as follows. Let

$$P_k(i+1) = \frac{\gamma_k^*}{\gamma_k(i)} P_k(i) \tag{14.15}$$

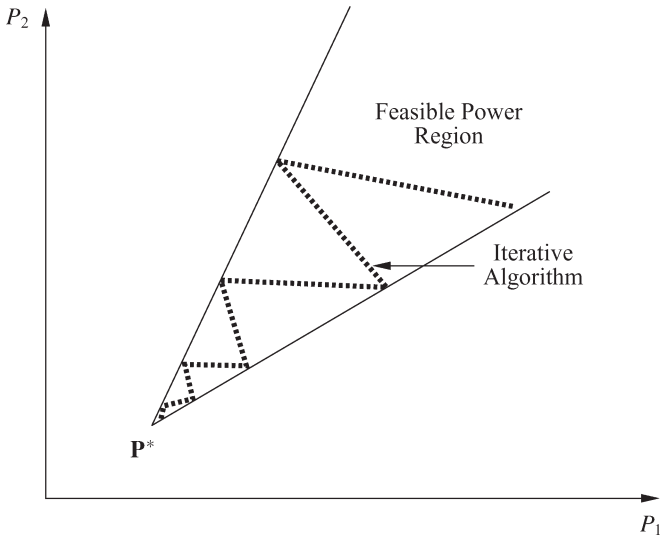


Figure 14.11: Iterative Foschini–Miljanic algorithm.

for each link $k \in \{1, 2, \ldots, N\}$. Hence, each transmitter increases power when its SINR is below its target and decreases power when its SINR exceeds its target. Measurements of SINR (or a function of them, such as BER) are typically made at the base station or access points, and a simple "up" or "down" command regarding transmit power can be fed back to each of the transmitters to perform the iterations. It is easy to show that (14.14) and (14.15) are pathwise equivalent and hence the per-user version of the power control algorithm also converges to $\mathbf{P}^*$. The feasible region of power vectors that achieve the SINR targets for a two-user system – along with the iterative algorithm that converges to the minimum power vector in this region – is illustrated in Figure 14.13. We see in this figure that the feasible region consists of all power pairs $\mathbf{P} = (P_1, P_2)$ that achieve a given pair of SINR targets and that the optimal pair $\mathbf{P}^*$ is the minimum power vector in this two-dimensional region.

The Foschini–Miljanic power control algorithm can also be combined with access control [37]. In this combination, access to the system is based on whether or not the new user causes other users to fall below their SINR targets. Specifically, when a new user requests access to the system, the base station or access point determines if a set of transmit powers exists such that he can be admitted without degrading existing users below their desired SINR threshold. If the new user cannot be accommodated in the system without violating the SINR requirements of existing users, then access is denied. If the new user can be accommodated then his power control algorithm and those of existing users are set to a feasible power vector under which all users (new and existing) meet their SINR targets.

A power control strategy for multiple access that takes into account delay constraints is proposed and analyzed in [38]. This strategy optimizes the transmit power relative to channel conditions and delay constraints via dynamic programming. The optimal strategy exhibits three modes: very low transmit power when the channel is poor and the tolerable delay large, higher transmit power when the channel and delay are average, and very high transmit power when the delay constraint is tight. This strategy exhibits significant power savings over constant transmit power while meeting the delay constraints of the traffic.

14.6 Downlink (Broadcast) Channel Capacity

When multiple users share the same channel, the channel capacity can no longer be characterized by a single number. At the extreme, if only one user occupies all signaling dimensions in the channel then the region reduces

to the single-user capacity described in Chapter 4. However, since there is an infinite number of ways to divide the channel between many users, the multiuser channel capacity is characterized by a *rate region*, where each point in the region is a vector of achievable rates that can be maintained by all the users simultaneously with arbitrarily small error probability. The union of achievable rate vectors under all multiuser transmission strategies is called the *capacity region* of the multiuser system. The channel capacity is different for uplink channels and downlink channels because of the fundamental differences between these channel models. However, the fact that downlink and uplink channels seem to be mirror images implies that there might be a connection between their capacities. In fact, there is a duality between these channels that allows the capacity region of either channel to be obtained from the capacity region of the other. Note that, in the analysis of channel capacity, the downlink is commonly referred to as the broadcast channel (BC) and the uplink is commonly referred to as the multiple access channel (MAC[5]), and we will use this terminology in our capacity discussions. In this section we describe the capacity region of the BC, Section 14.7 treats the MAC capacity region, and Section 14.8 characterizes the duality between these two channels and how it can be exploited in capacity calculations.

After first describing the AWGN BC model, we will characterize its rate region using superposition code division (CD) with successive interference cancellation, time division (TD), and frequency division (FD). We then obtain the rate regions using DSSS for orthogonal and nonorthogonal codes. The BC and corresponding capacity results under fading are also treated.

We will see that capacity is achieved using superposition CD with interference cancellation. In addition, DSSS with successive interference cancellation has a capacity penalty (relative to superposition coding) that increases with spreading gain. Finally, spread spectrum with orthogonal CD can achieve a subset of the TD and FD capacity regions, but spread spectrum with nonorthogonal coding and no interference cancellation is inferior to all the other spectrum-sharing techniques. The capacity regions in fading depend on what is known about the fading channel at the transmitter and receiver, which is analogous to the case of single-user capacity in fading.

14.6.1 Channel Model

We consider a broadcast channel consisting of one transmitter sending different data streams also called independent information or data – to different receivers. Thus, our model is not applicable to a typical radio or TV broadcast channel, where the same data stream – also called common information or data – is received by all users. However, the capacity results easily extend to include common data, as described in Section 14.6.3. The capacity region of the BC characterizes the rates at which information can be conveyed to the different receivers simultaneously. We mainly focus on capacity regions for the two-user BC, since the general properties and the relative performance of the different spectrum-sharing techniques are the same for any finite number of users [39].

The two-user BC has one transmitter and two distant receivers receiving data at rate R_k, $k = 1, 2$. The channel power gain between the transmitter and kth receiver is g_k, $k = 1, 2$, and each receiver has AWGN of power spectral density (PSD) $N_0/2$. We define the effective noise on the kth channel as $n_k = N_0/g_k$, $k = 1, 2$, and we arbitrarily assume that $n_1 \leq n_2$; that is, we assume the first user has a larger channel gain to its receiver than the second user. Incorporating the channel gains into the noise PSD does not change the SINR for any user, since the signal and interference on each user's channel are attenuated by the same channel gain. Thus, the BC capacity with channel gains $\{g_k\}$ is the same as the BC capacity based on the effective noises $\{n_k\}$ [40]. The fact that the channel gains (or, equivalently, the effective noise of the users) can be ordered makes the channel model a *degraded broadcast channel*, for which a general formula for channel capacity is known [41, Chap. 14.6]. We denote the transmitter's total average power and bandwidth by P and B, respectively.

If the transmitter allocates all the power and bandwidth to one of the users, then clearly the other user will receive a rate of zero. Therefore, the set of simultaneously achievable rates (R_1, R_2) includes the pairs $(C_1, 0)$ and

[5]MAC is also used as an abbreviation for the "medium access control" layer in networks [1, Chap. 1.2].

$(0, C_2)$, where

$$C_k = B \log_2 \left(1 + \frac{P}{n_k B}\right), \quad k = 1, 2, \tag{14.16}$$

is the single-user capacity in bits per second for an AWGN channel, as given in Section 4.1. These two points bound the BC capacity region. We now consider rate pairs in the interior of the region that are achieved using more equitable methods of dividing the system resources.

14.6.2 Capacity in AWGN

In this section we compute the set of achievable rate vectors of the AWGN BC under TD, FD, and the optimal method of superposition coding, which achieves capacity. In time division, the transmit power P and bandwidth B are allocated to user 1 for a fraction τ of the total transmission time and then to user 2 for the remainder of the transmission. This TD scheme achieves a straight line between the points C_1 and C_2, corresponding to the rate pairs

$$C_{\text{TD}} = \bigcup_{\{\tau: 0 \le \tau \le 1\}} \left(R_1 = \tau B \log_2 \left(1 + \frac{P}{n_1 B}\right), R_2 = (1-\tau) B \log_2 \left(1 + \frac{P}{n_2 B}\right)\right). \tag{14.17}$$

This equal-power TD achievable rate region is illustrated in Figures 14.13 and 14.14 (pp. 489–490), where $n_1 B$ and $n_2 B$ differ by 3 dB and 20 dB, respectively. This dB difference, which reflects the difference in the channel gains of the two users, is a crucial parameter in comparing the achievable rates of the different spectrum-sharing techniques, as we discuss in more detail below.

If we also vary the transmit power of each user, subject to an average power constraint P, then we can obtain a larger set of achievable rates. Let P_1 and P_2 denote the power allocated to users 1 and 2, respectively, over their assigned time slots. The average power constraint then becomes $\tau P_1 + (1-\tau) P_2 = P$. The achievable rate region with TD and variable power (VP) allocation is then

$$\mathcal{C}_{\text{TD,VP}} = \bigcup_{\{\tau, P_1, P_2: 0 \le \tau \le 1; \tau P_1 + (1-\tau) P_2 = P\}} \left(R_1 = \tau B \log_2 \left(1 + \frac{P_1}{n_1 B}\right), \right.$$
$$\left. R_2 = (1-\tau) B \log_2 \left(1 + \frac{P_2}{n_2 B}\right)\right). \tag{14.18}$$

In frequency division the transmitter allocates P_k of its total power P and B_k of its total bandwidth B to user k. The power and bandwidth constraints require that $P_1 + P_2 = P$ and $B_1 + B_2 = B$. The set of achievable rates for a fixed frequency division (B_1, B_2) is thus

$$\mathcal{C}_{\text{FFD}} = \bigcup_{\{P_1, P_2: P_1 + P_2 = P\}} \left(R_1 = B_1 \log_2 \left(1 + \frac{P_1}{n_1 B_1}\right), R_2 = B_2 \log_2 \left(1 + \frac{P_2}{n_2 B_2}\right)\right). \tag{14.19}$$

It was shown by Bergmans and Cover [39] that, for n_1 strictly less than n_2 and for any fixed frequency division (B_1, B_2), there exists a range of power allocations $\{P_1, P_2 \colon P_1 + P_2 = P\}$ whose corresponding rate pairs exceed a segment of the equal-power TD line (14.17). This superiority is also illustrated in Figures 14.13 and 14.14, where the rate regions for fixed FD under two different bandwidth divisions are plotted. The superiority is difficult

to distinguish in Figure 14.13, where the users have similar channel gains, but is much more apparent in Figure 14.14, where the users have a 20-dB difference in gain.

The FD achievable rate region is defined as the union of fixed FD rate regions (14.19) over all bandwidth divisions:

$$\mathcal{C}_{\mathrm{FD}} = \bigcup_{\{P_1,P_2,B_1,B_2:P_1+P_2=P;B_1+B_2=B\}} \left(R_1 = B_1 \log_2 \left(1 + \frac{P_1}{n_1 B_1} \right), \right.$$
$$\left. R_2 = B_2 \log_2 \left(1 + \frac{P_2}{n_2 B_2} \right) \right). \quad (14.20)$$

It was shown in [39] that this achievable rate region exceeds the equal-power TD rate region (14.17). This superiority is indicated by the closure of the fixed FD regions in Figures 14.13 and 14.14 – although it is difficult to see in Figure 14.13, where the users have a similar received signal-to-noise ratio. In fact, when $n_1 = n_2$, (14.20) reduces to (14.17) [39]. Thus, optimal power and/or frequency allocation is more beneficial when the users have more disparate channel quality.

Note that the achievable rate region for TD with unequal power allocation given by (14.18) is the same as the FD achievable rate region (14.20). This is seen by letting $B_i = \tau_i B$ and $\pi_i = \tau_i P_i$ in (14.18), where $\tau_1 = \tau$ and $\tau_2 = 1 - \tau$. The power constraint then becomes $\pi_1 + \pi_2 = P$. Making these substitutions in (14.18) yields

$$\mathcal{C}_{\mathrm{TD,VP}} = \bigcup_{\{\pi_1,\pi_2:\pi_1+\pi_2=P\}} \left(R_1 = B_1 \log_2 \left(1 + \frac{\pi_1}{n_1 B_1} \right), R_2 = B_2 \log_2 \left(1 + \frac{\pi_2}{n_2 B_2} \right) \right). \quad (14.21)$$

By comparing this with (14.19) we see that, with appropriate choice of P_k and τ_k, any point in the FD achievable rate region can also be achieved through TD with variable power.

Superposition coding with successive interference cancellation is a multiresolution coding technique whereby the user with the higher channel gain can distinguish the fine resolution of the received signal constellation while the user with the lower channel gain can only distinguish the constellation's coarse resolution (see [39] and [41, Chap. 14.6]). An example of a two-level superposition code constellation taken from [42] is 32-QAM with embedded 4-PSK, as shown in Figure 14.12. In this example, the transmitted constellation point is one of the 32-QAM signal points chosen as follows. The data stream intended for the user with the worse SNR (user 2 in our model, since $n_2 > n_1$) provides 2 bits to select one of the 4-PSK superpoints. The data stream intended for the user with the better SNR provides 3 bits to select one of the eight constellation points surrounding the selected superpoint. After transmission through the channel, the user with the better SNR can easily distinguish the quadrant in which the constellation point lies. Thus, the 4-PSK superpoint is effectively subtracted out by this user. However, the user with the worse SNR cannot distinguish between the 32-QAM points around its 4-PSK superpoints. Hence, the 32-QAM modulation superimposed on the 4-PSK modulation appears as noise to this user, and this user can only decode the 4-PSK. These ideas can be easily extended to multiple users using more complex signal constellations. Since superposition coding achieves multiple rates by expanding its signal constellation, it does not require bandwidth expansion.

The two-user capacity region using superposition coding and successive interference cancellation was derived in [39] to be the set of rate pairs

$$\mathcal{C}_{\mathrm{BC}} = \bigcup_{\{P_1,P_2:P_1+P_2=P\}} \left(R_1 = B \log_2 \left(1 + \frac{P_1}{n_1 B} \right), R_2 = B \log_2 \left(1 + \frac{P_2}{n_2 B + P_1} \right) \right). \quad (14.22)$$

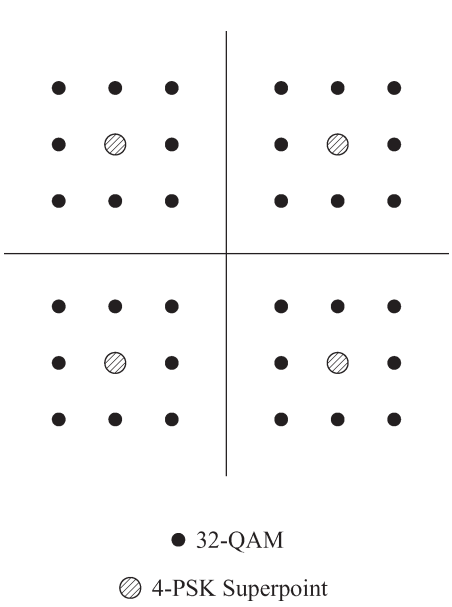


Figure 14.12: 32-QAM with embedded 4-PSK.

The intuitive explanation for (14.22) is the same as for the example illustrated in Figure 14.12. Since $n_1 < n_2$, user 1 correctly receives all the data transmitted to user 2. Therefore, user 1 can decode and subtract out user 2's message and then decode its own message. User 2 cannot decode the message intended for user 1, since it has a worse SNR. Thus, user 1's message, with power P_1, contributes an additional noise term to user 2's received message. This message can be treated as an additional AWGN term because the capacity-achieving distributions for the signals associated with each user are Gaussian (see [39] and [41, Chap. 14.1]). This same process is used by the successive interference cancellation method for DSSS described in Section 13.4.4. However, although successive interference cancellation achieves the capacity region (14.22), it is not necessarily the best method to use in practice. The capacity analysis assumes perfect signal decoding and channel estimation, whereas real systems exhibit some decoding and channel estimation error. This error leads to decision-feedback errors in the successive interference cancellation scheme. Thus, multiuser detection methods that do not suffer from this type of error may work better in practice than successive cancellation does.

The rate region defined by (14.22) was shown in [43] to exceed the regions achievable through either TD or FD when $n_1 < n_2$. Moreover, it was also shown in [43] that this is the maximum achievable set of rate pairs for any type of coding and spectrum sharing and thus (14.22) defines the BC capacity region; hence the notation C_{BC}. However, if all users have the same SNR, then this capacity region collapses to the equal-power TD line (14.17). Thus, when $n_1 = n_2$, all the spectrum-sharing methods have the same rate region.

The ideas of superposition coding are easily extended to a K-user system for $K > 2$. Assume a BC with K users, each with channel gain g_k. We first order the users relative to their effective noise $n_k = N_0/g_k$. Based on this effective noise ordering, the superposition coding will now have K levels, where the coarsest level can be detected by the user with the largest effective noise (worst SNR), the next level can be detected by the user with the next largest effective noise, and so forth. Each user can remove the effects of the constellation points associated with other users who have noisier channels, but the constellation points transmitted to users with better channels appear as additional noise. Assuming a total power constraint P, the multiuser extension to the two-user region (14.22) is given by

$$\mathcal{C}_{\mathrm{BC}} = \bigcup_{\{P_k : \Sigma_{k=1}^K P_k = P\}} \left\{ (R_1, \ldots, R_K) \colon R_k = B \log_2 \left(1 + \frac{P_k}{n_k B + \sum_{j=1}^K P_j \mathbf{1}[n_k > n_j]} \right) \right\}, \tag{14.23}$$

where $\mathbf{1}[\cdot]$ denotes the indicator function.

We define the *sum-rate* (SR) *capacity* of a BC as the maximum sum of rates taken over all rate vectors in the capacity region:

$$C_{\mathrm{BCSR}} = \max_{(R_1,\ldots,R_K) \in C_{\mathrm{BC}}} \sum_{k=1}^K R_k. \tag{14.24}$$

Sum-rate capacity is a single number that defines the maximum throughput of the system (regardless of fairness in the rate allocation between the users). It is therefore much easier to characterize than the K-dimensional capacity region and often leads to important insights. In particular, it can be shown from (14.23) that sum-rate capacity is achieved on the AWGN BC by assigning all power P to the user with the highest channel gain or, equivalently, the lowest effective noise. Defining $n_{\min} \triangleq \min_k n_k$ and $g_{\max} \triangleq \max_k g_k$, this implies that the sum-rate capacity C_{BCSR} for the K-user AWGN BC is given by

$$C_{\mathrm{BCSR}} = B \log_2 \left(1 + \frac{P}{n_{\min} B} \right) = B \log_2 \left(1 + \frac{g_{\max} P}{N_0 B} \right). \tag{14.25}$$

The sum-rate point is therefore one of the boundary points (14.16) of the capacity region, which is the same for superposition coding, TD, and FD because all resources are assigned to a single user.

Example 14.5: Consider an AWGN BC with total transmit power $P = 10$ mW, $n_1 = 10^{-9}$ W/Hz, $n_2 = 10^{-8}$ W/Hz, and $B = 100$ kHz. Suppose user 1 requires a data rate of 300 kbps. Find the rate that can be allocated to user 2 under equal-power TD, equal-bandwidth FD, and superposition coding.

Solution: In equal-power TD, user 1 has a rate of $R_1 = \tau B \log_2(1 + P/n_1 B) = (6.658 \cdot 10^5)\tau$ bps. Setting R_1 to the desired value $R_1 = (6.658 \cdot 10^5)\tau = 3 \cdot 10^5$ bps and solving for τ yields $\tau = 3 \cdot 10^5 / 6.644 \cdot 10^5 = .451$. Then user 2 gets a rate of $R_2 = (1 - \tau) B \log_2(1 + P/n_2 B) = 1.89 \cdot 10^5$ bps. In equal-bandwidth frequency division we require $R_1 = .5B \log_2(1 + P_1/.5 n_1 B) = 3 \cdot 10^5$ bps. Solving for $P_1 = .5 n_1 B (2^{R_1/(.5B)} - 1)$ yields $P_1 = 3.15$ mW. Setting $P_2 = P - P_1 = 6.85$ mW, we get $R_2 = .5B \log_2(1 + P_2/.5 n_2 B) = 1.94 \cdot 10^5$ bps. Finally, with superposition coding we have $R_1 = B \log_2(1 + P_1/n_1 B) = 3 \cdot 10^5$. Solving for $P_1 = n_1 B (2^{R_1/B} - 1)$ yields $P_1 = .7$ mW. Then

$$R_2 = B \log \left(1 + \frac{P - P_1}{n_2 B + P_1} \right) = 2.69 \cdot 10^5 \text{ bps}. \tag{14.26}$$

Superposition coding is clearly superior to both TD and FD, as expected, although the performance of these techniques would be closer to that of superposition coding if we optimized the power allocation for TD or the bandwidth allocation for FD.

Example 14.6: Find the sum-rate capacity for the system in the prior example.

Solution: We have $P = 10$ mW, $n_1 = 10^{-9}$ W/Hz, $n_2 = 10^{-8}$ W/Hz, and $B = 100$ kHz. The minimum noise is associated with user 1, $n_{\min} = 10^{-9}$. Thus, $C_{\mathrm{BCSR}} = B \log_2(1 + P/n_{\min} B) = 6.644 \cdot 10^5$ bps. This sum-rate is achievable with TD, FD, or superposition coding, which are all equivalent for this sum-rate capacity because all

resources are allocated to the first user.

Code division for multiple users can also be implemented using DSSS, as discussed in Section 13.4. In such systems the modulated data signal for each user is modulated by a unique spreading code, which increases the transmit signal bandwidth by approximately G, the processing gain of the spreading code. For orthogonal spreading codes, the cross-correlation between the respective codes is zero, and these codes require a spreading gain of N to produce N orthogonal codes. For a total bandwidth constraint B, the information bandwidth of each user's signal with these spreading codes is thus limited to B/N. The two-user achievable rate region with these orthogonal codes (OC) is then

$$\mathcal{C}_{\text{DS,OC}} = \bigcup_{\{P_1,P_2:P_1+P_2=P\}} \left(R_1 = \frac{B}{2}\log_2\left(1+\frac{P_1}{n_1 B/2}\right), R_2 = \frac{B}{2}\log_2\left(1+\frac{P_2}{n_2 B/2}\right)\right). \tag{14.27}$$

Comparing (14.27) with (14.19), we see that DSSS with orthogonal coding has the same capacity region as fixed FD with the bandwidth equally divided ($B_1 = B_2 = B/2$). From (14.21), TD with unequal power allocation can also achieve all points in this rate region. Thus, DSSS with orthogonal (e.g., Walsh–Hadamard) codes achieves a subset of the TD and FD achievable rate regions. More general orthogonal codes are needed to achieve the same region as these other techniques.

We now consider DSSS with nonorthogonal spreading codes. As discussed in Section 13.4.2, in these systems interference between users is attenuated by the code cross-correlation. Thus, if interference is treated as noise, its power contribution to the SIR is reduced by the square of the code cross-correlation. From (13.6), we will assume that spreading codes with a processing gain of G reduce the interference power by $1/G$. This is a reasonable approximation for random spreading codes, although (as discussed in Section 13.4.1) the exact value of the interference power reduction depends on the nature of the spreading codes and other assumptions [44, 45]. Since the signal bandwidth is increased by G, the two-user BC rate region achievable through DSSS with nonorthogonal codes (NC) and successive interference cancellation (IC) is given by

$$\mathcal{C}_{\text{DS,NC,IC}} \bigcup_{\{P_1,P_2:P_1+P_2=P\}} \left(R_1 = \frac{B}{G}\log_2\left(1+\frac{P_1}{n_1 B/G}\right), \right.$$
$$\left. R_2 = \frac{B}{G}\log_2\left(1+\frac{P_2}{n_2 B/G + P_1/G}\right)\right). \tag{14.28}$$

By the concavity of the log function, the rate region defined by (14.28) for $G > 1$ is smaller than the rate region (14.22) obtained using superposition coding, and the degradation increases with increasing values of G. This implies that for DSSS with nonorthogonal coding, the spreading gain should be minimized in order to maximize capacity.

With nonorthogonal coding and no interference cancellation, the receiver treats all signals intended for other users as noise, resulting in the achievable rate region

$$\mathcal{C}_{\text{DS, NC}} = \bigcup_{\{P_1,P_2:P_1+P_2=P\}} \left(R_1 = \frac{B}{G}\log_2\left(1+\frac{P_1}{n_1 B/G + P_2/G}\right), \right.$$
$$\left. R_2 = \frac{B}{G}\log_2\left(1+\frac{P_2}{n_2 B/G + P_1/G}\right)\right). \tag{14.29}$$

Again using concavity of the log function, we have that $G = 1$ maximizes this rate region and that the rate region decreases as G increases. Moreover, the radius of curvature for (14.29) is given by

$$\chi = \frac{\dot{R}_1 \ddot{R}_2 - \dot{R}_2 \ddot{R}_1}{(\dot{R}_1^2 + \dot{R}_2^2)^{3/2}}, \tag{14.30}$$

where $\dot{R}_i$ and $\ddot{R}_i$ denote (respectively) the first and second derivatives of R_i with respect to α for $P_1 = \alpha P$ and $P_2 = (1 - \alpha)P$. For $G = 1, \chi \geq 0$. Thus, the rate region for nonorthogonal coding without interference cancellation (14.29) is bounded by a convex function with end points C_1 and C_2, as shown in Figures 14.13 and 14.14. Therefore, the achievable rate region for DSSS with nonorthogonal coding and no interference cancellation will lie beneath the regions for TD and FD, which are bounded by concave functions with the same endpoints.

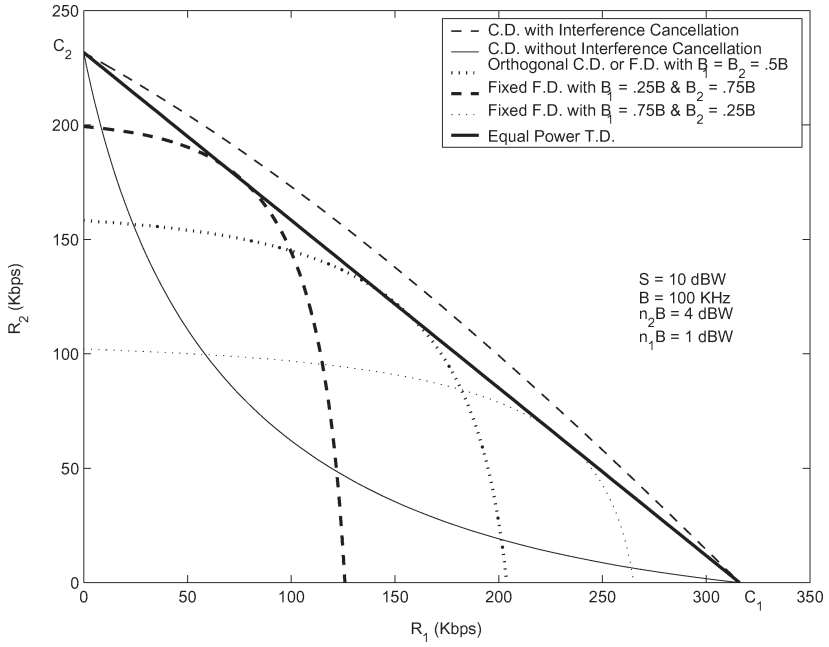


Figure 14.13: Two-user capacity region with 3-dB SNR difference.

The achievable rate regions for equal-power TD (14.17), FD (14.19), orthogonal DSSS (14.27), and nonorthogonal DSSS with (14.22) and without (14.29) interference cancellation are illustrated in Figures 14.13 and 14.14, where the SNR between the users differs by 3 dB and 20 dB, respectively. For the calculation of (14.29) we assume $G = 1$. Direct-sequence spread spectrum with larger values of the spreading gain will result in a smaller rate region.

14.6.3 Common Data

In many broadcasting applications, common data is sent to all users in the system. For example, television and radio stations broadcast the same data to all users, and in wireless Internet applications many users may want to download the same stock quotes and sports scores. The nature of superposition coding makes it straightforward to develop optimal broadcasting techniques for common data and to incorporate common data into the capacity region for the broadcast channel. In particular, for a two-user BC with superposition coding, the user with the better channel gain always receives the data intended for the user with the worse channel gain, along with his own data. Thus, since common data must be transmitted to both users, we can encode all common data as independent

data intended for the user with the worse SNR. Since the user with the better SNR will also receive this data, it will be received by both users.

Under this transmission strategy, if the rate pair (R_1, R_2) is in the capacity region of the two-user BC with independent data defined by (14.22), then for any $R_0 \leq R_2$ we can achieve the rate triple $(R_0, R_1, R_2 - R_0)$ for the BC with common and independent data, where R_0 is the rate of common data, R_1 is the rate of user 1's independent data, and $R_2 - R_0$ is the rate of user 2's independent data. Mathematically, this gives the three-dimensional capacity region

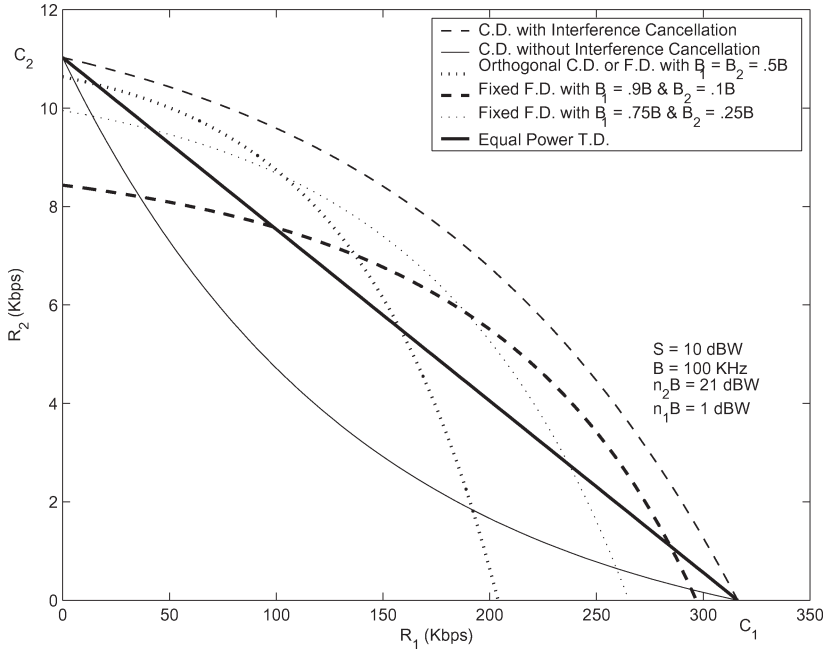


Figure 14.14: Two-user capacity region with 20-dB SNR difference.

$$\mathcal{C}_{\mathrm{BC}} = \bigcup_{\{P_1, P_2 : P_1 + P_2 = P\}} \left(R_0 \leq B \log_2 \left(1 + \frac{P_2}{n_2 B + P_1} \right), R_1 = B \log_2 \left(1 + \frac{P_1}{n_1 B} \right), \right.$$
$$\left. R_2 = B \log_2 \left(1 + \frac{P_2}{n_2 B + P_1} \right) - R_0 \right). \tag{14.31}$$

Example 14.7: In Example 14.5, we saw that for a broadcast channel with total transmit power $P = 10\,\mathrm{mW}, n_1 = 10^{-9}\ \mathrm{W/Hz}, n_2 = 10^{-8}$ W/Hz, and $B = 100$ kHz, the rate pair $(R_1, R_2) = (3 \cdot 10^5, 2.69 \cdot 10^5)$ is on the boundary of the capacity region. Suppose that user 1 desires an independent data rate of 300 kbps and that a common data rate of 100 kbps is required for the system. At what rate can user 2 get independent data?

Solution: In order for $R_1 = 300$ kbps, we require the same $P_1 = .7$ mW as in Example 14.5.2. The common information rate $R_0 = 10^5 < 2.69 \cdot 10^5$, so from (14.31) it follows that the independent information rate to user 2 is just $R_2 - R_0 = 2.69 \cdot 10^5 - 1 \cdot 10^5 = 1.69 \cdot 10^5$ bps.

14.6.4 Capacity in Fading

We now consider the capacity region of BCs with fading, where the users have independent random channel gains that change over time. As described in Section 4.2 for single-user channels, the capacity of fading broadcast channels depends on what is known about the channel at the transmitter and receiver. However, capacity of a BC is only known for degraded BCs, and this model requires that the channel gains are known to both the transmitter and receiver. Furthermore, superposition coding cannot be used without transmitter knowledge of the channel gains, since if the transmitter does not know the relative channel gains then it does not know which user can receive the coarse constellation point and which can receive the fine one. Thus we will only consider fading BCs where there is perfect channel side information (CSI) about the instantaneous channel gains at the transmitter and at all receivers. We also assume that the channel is slowly fading so that, for a given fading state, the coding strategy that achieves any point in the capacity region for the static BC with this state has sufficient time to drive the error probability close to zero before the channel gains change.[6]

As with the single-user fading channel, there are two notions of capacity for multiuser fading channels with perfect CSI: ergodic (Shannon) capacity and outage capacity. The ergodic capacity region of a multiuser fading channel characterizes the achievable rate vectors averaged over all fading states [40, 46] while the outage capacity region dictates the set of fixed-rate vectors that can be maintained in all fading states subject to a given outage probability [47, 48, 49]. Zero-outage capacity refers to outage capacity with zero outage probability [47] – that is, the set of fixed-rate vectors that can be maintained in all fading states. The ergodic capacity region, which is analogous to ergodic capacity for single-user systems, defines the data-rate vectors that can be maintained over time without any constraints on delay. Hence, in some fading states the data rate may be small or zero, which can be problematic for delay-constrained applications like voice or video. The outage capacity region – analogous to outage capacity in single-user systems – forces a fixed-rate vector in all non-outage fading states, which is perhaps a more appropriate capacity metric for delay-constrained applications. However, the requirement for maintaining a fixed rate even in very deep fades can severely decrease the outage capacity region relative to the ergodic capacity region. In fact, the zero-outage capacity region when all users exhibit Rayleigh fading is zero for all users.

We consider a BC with AWGN and fading where a single transmitter communicates independent information to K users over bandwidth B with average transmit power $\bar{P}$. The transmitter and all receivers have a single antenna. The time-varying power gain of user k's channel at time i is $g_k[i]$. Each receiver has AWGN with PSD $N_0/2$. We define the effective time-varying noise of the kth user as[7] $n_k[i] = N_0/g_k[i]$. The *effective noise vector* at time i is defined as

$$\mathbf{n}[i] = (n_1[i], \ldots, n_K[i]). \tag{14.32}$$

We also call this the *fading state* at time i, since it characterizes the channel gains $g_k[i]$ associated with each user at time i. We will denote the kth element of this vector as $n_k[i]$ or just n_k when the time reference is clear. As with the static channel, the capacity of the fading BC can be computed based on its time-varying channel gains or its time-varying effective noise vector. The ergodic BC capacity region is defined as the set of all average rates achievable in a fading channel with arbitrarily small probability of error, where the average is taken with respect to all fading states. In [40], the ergodic capacity region and optimal power allocation scheme for the fading BC is found by decomposing the fading channel into a parallel set of static BCs, one for every possible fading state $\mathbf{n} = (N_0/g_1, \ldots, N_0/g_K)$. In each fading state the channel can be viewed as a static AWGN BC, and time, frequency, or code division can be applied to the channel in each fading state.

[6] More precisely, the coding strategy that achieves a point in the AWGN BC capacity region uses a block code, and the error probability of the code goes to zero with blocklength. Our slow fading assumption presumes that the channel gains stay constant long enough for the block code associated with these gains to drive the error probability close to zero.

[7] Notice that the noise vector is the instantaneous power of the noise and not the instantaneous noise sample.

Since the transmitter and all receivers know $\mathbf{n}[i]$, superposition coding according to the ordering of the current effective noise vector can be used by the transmitter. Each receiver can perform successive decoding in which the users with larger effective noise are decoded and subtracted out before decoding the desired signal. Furthermore, the power transmitted to each user $P_j(\mathbf{n})$ is a function of the current fading state. Since the transmission scheme is based on superposition coding, it remains only to determine the optimal power allocation across users and over time.

We define a power policy $\mathcal{P}$ over all possible fading states as a function that maps from any fading state $\mathbf{n}$ to the transmitted power $P_k(\mathbf{n})$ for each user. Let $\mathcal{F}_{\mathrm{BC}}$ denote the set of all power policies satisfying average power constraint $\bar{P}$:

$$\mathcal{F}_{\mathrm{BC}} \stackrel{\triangle}{=} \left\{ \mathcal{P} : \mathbf{E}_{\mathbf{n}} \left[\sum_{k=1}^{K} P_k(\mathbf{n}) \right] \leq \bar{P} \right\}. \tag{14.33}$$

From (14.23), the capacity region assuming a constant fading state $\mathbf{n}$ with power allocation $\mathbf{P}(\mathbf{n}) = \{P_k(\mathbf{n}) : k = 1, \ldots, K\}$ is given by

$$\mathcal{C}_{\mathrm{BC}}(\mathbf{P}(\mathbf{n})) = \Bigg\{ (R_1(\mathbf{P}(\mathbf{n})), \ldots, R_K(\mathbf{P}(\mathbf{n})) : R_k(\mathbf{P}(\mathbf{n})) = B \log_2 \left(1 + \frac{P_k(\mathbf{n})}{n_k B + \sum_{j=1}^{K} P_j(\mathbf{n}) \mathbf{1}[n_k > n_j]} \right) \Bigg\}. \tag{14.34}$$

Let $\mathcal{C}_{\mathrm{BC}}(\mathcal{P})$ denote the set of achievable rates averaged over all fading states for power policy $\mathcal{P}$:

$$\mathcal{C}_{\mathrm{BC}}(\mathcal{P}) = \{(R_1, \ldots, R_K) : R_k \leq \mathbf{E}_{\mathbf{n}}[R_k(\mathbf{P}(\mathbf{n}))]\}, \tag{14.35}$$

where $R_k(\mathbf{P}(\mathbf{n}))$ is as given in (14.34). From [40], the ergodic capacity region of the BC with perfect CSI and power constraint $\bar{P}$ is:

$$\mathcal{C}_{\mathrm{BC}}(\bar{P}) = \bigcup_{\mathcal{P} \in \mathcal{F}_{\mathrm{BC}}} \mathcal{C}_{\mathrm{BC}}(\mathcal{P}). \tag{14.36}$$

It is further shown in [40] that the region $\mathcal{C}_{\mathrm{BC}}(\bar{P})$ is convex and that the optimal power allocation scheme is an extension of water-filling with K different water levels for a K user system.

We can also define achievable rate vectors for TD or FD, although these will clearly lie inside the ergodic capacity region, since superposition coding outperforms both of these techniques in every fading state. The optimal form of TD adapts the power assigned to each user relative to the current fading state. Similarly, the optimal form of FD adapts the bandwidth and power assigned to each user relative to the current fading state. As described in Section 14.6.2, for each fading state, varying the power in TD yields the same rates as varying the power and bandwidth in FD. Thus, the achievable rates for these two techniques averaged over all fading states are the same. Focusing on the FD region, assume a power policy $\mathcal{P} \in \mathcal{F}_{\mathrm{BC}}$ that assigns power $P_k(\mathbf{n})$ to the kth user in fading state $\mathbf{n}$. By (14.33), a power policy $\mathcal{P} \in \mathcal{F}_{\mathrm{BC}}$ satisfies the average power constraint. Also assume a bandwidth policy B that assigns bandwidth $B_k(\mathbf{n})$ to user k in state $\mathbf{n}$, and let $\mathcal{G}$ denote the set of all bandwidth policies satisfying the bandwidth constraint of the system:

$$\mathcal{G} \triangleq \left\{ \mathcal{B} : \sum_{k=1}^{K} B_k(\mathbf{n}) = B, \ \forall \mathbf{n} \right\}. \tag{14.37}$$

The set of achievable rates for FD under policies $\mathcal{P}$ and $\mathcal{B}$ is

$$\mathcal{C}_{\mathrm{FD}}(\mathcal{P},\mathcal{B}) = \{(R_1,\ldots,R_K) : R_k \le \mathbf{E}_{\mathbf{n}}[R_k(\mathbf{P}(\mathbf{n}),\mathcal{B})]\}, \tag{14.38}$$

where

$$R_k(P(\mathbf{n}),\mathcal{B}) = B_k(\mathbf{n})\log_2\left(1+\frac{P_k(\mathbf{n})}{n_k B_k(\mathbf{n})}\right). \tag{14.39}$$

The set of all achievable rates under frequency division with perfect CSI, subject to power constraint $\bar{P}$ and bandwidth constraint B, is then

$$\mathcal{C}_{\mathrm{FD}}(\bar{P},B) = \bigcup_{\mathcal{P}\in\mathcal{F}_{\mathrm{BC}},B\in\mathcal{G}} \mathcal{C}_{\mathrm{FD}}(\mathcal{P},\mathcal{B}). \tag{14.40}$$

The sum-rate capacity for fading BCs is defined as the maximum sum of achievable rates, maximized over all rate vectors in the ergodic BC capacity region. Since sum-rate capacity for the AWGN BC is maximized by transmitting only to the user with the best channel, in fading the sum-rate is maximized by transmitting only to the user with the best channel in each channel state. Clearly superposition CD, TD, and FD are all equivalent in this setting, since all resources are assigned to a single user in each state. We can compute the sum-rate capacity and the optimal power allocation over time from an equivalent single-user fading channel with time-varying effective noise $n[i] = \min_k n_k[i]$ and average power constraint $\bar{P}$. From Section 4.2.4 the optimal power allocation to the user with the best channel at time i is thus a water-filling in time, with cutoff value determined from the distribution of $\min_k n_k[i]$.

The ergodic capacity and achievable rate regions for fading broadcast channels under CD, TD, and FD are computed in [40] for different fading distributions, along with the optimal adaptive resource allocation strategies that achieve the boundaries of these regions. These adaptive transmission policies exploit *multiuser diversity* in that more resources (power, bandwidth, timeslots) are allocated to the users with the best channels in any given fading state. In particular, sum-rate capacity is achieved by allocating all resources in any given state to the user with the best channel. Multiuser diversity will be discussed in more detail in Section 14.9.

The zero-outage BC capacity region defines the set of rates that can be simultaneously achieved for all users in *all* fading states while meeting the average power constraint; it is the multiuser extension of zero-outage capacity defined in Section 4.2.4 for single-user channels. From [48], the power required to support a rate vector $\mathbf{R} = (R_1,R_2,\ldots,R_K)$ in fading state $\mathbf{n}$ is:

$$P^{\min}(\mathbf{R},\mathbf{n}) = \sum_{k=1}^{K-1}\left[2^{\left(\sum_{j=k+1}^{K} R_{\pi(j)}/B\right)}(2^{R_{\pi(k)}/B}-1)n_{\pi(k)}B\right] + \left(2^{R_{\pi(K)}/B}-1\right)n_{\pi(K)}B, \tag{14.41}$$

where $\pi(\cdot)$ is the permutation such that

$$n_{\pi(1)} < n_{\pi(2)} < \ldots < n_{\pi(K)}. \tag{14.42}$$

Therefore, the zero-outage capacity region is the union of all rate vectors that meet the average power constraint:

$$\mathcal{C}^0_{\mathrm{BC}}(\bar{P}) = \bigcup_{\{\mathbf{R}:\mathbf{E}_{\mathbf{n}}[P^{\min}(\mathbf{R},\mathbf{n})]\le\bar{P}\}} \mathbf{R} = (R_1,R_2,\ldots,R_K). \tag{14.43}$$

The boundary of the zero-outage capacity region is the set of all rate vectors $\mathbf{R}$ such that the power constraint is met with equality. For the two-user BC with time-varying AWGN with powers n_1 and n_2, this boundary simplifies to the set of all (R_1,R_2) that satisfy the following equation [48]:

$$\bar{P} = p(n_1 < n_2)[\mathbf{E}[n_1 B | n_1 < n_2] 2^{R_2/B}(2^{R_1/B} - 1) + \mathbf{E}[n_2 B | n_1 < n_2](2^{R_2/B} - 1)]$$
$$+ p(n_1 \geq n_2)[\mathbf{E}[n_2 B | n_1 \geq n_2] 2^{R_1/B}(2^{R_2/B} - 1) + \mathbf{E}[n_1 B | n_1 \geq n_2](2^{R_1/B} - 1)]. \quad (14.44)$$

The boundary is determined solely by $\mathbf{E}[n_1 B | n_1 < n_2], \mathbf{E}[n_2 B | n_1 < n_2], \mathbf{E}[n_1 B | n_1 \geq n_2]$, and $\mathbf{E}[n_2 B | n_1 \geq n_2]$. This follows because the power required to achieve a rate vector is a linear function of the noise levels in each state, as seen in (14.41). The zero-outage capacity region depends on the conditional expectations of the noises – as opposed to their unconditional expectations – since every different ordering of noises leads to a different expression for the required power in each state, as can be seen from (14.41).

The outage capacity region of the BC is defined similarly as the zero-outage capacity region, except that users may have some nonzero probability of outage so that they can suspend transmission in some outage states. This provides additional flexibility in the system because, under severe fading conditions, maintaining a fixed rate in *all* fading states can consume a great deal of power. In particular, we saw in Section 4.2.4 that, for a single-user fading channel, maintaining any nonzero fixed rate in Rayleigh fading requires infinite power. By allowing some outage, power can be conserved from outage states to maintain higher rates in non-outage states. The outage capacity region is more difficult to obtain than the zero-outage capacity region, since in any given fading state the transmission strategy must determine which users to put into outage. Once the outage users are determined, the power required to maintain the remaining users is given by (14.41) for the rate vector associated with the $K' \leq K$ users that are not in outage. It is shown in [48] that this decision should be made based on a threshold policy, and the resulting outage capacity region is then obtained implicitly based on the threshold policy and the power allocation (14.41) for non-outage users.

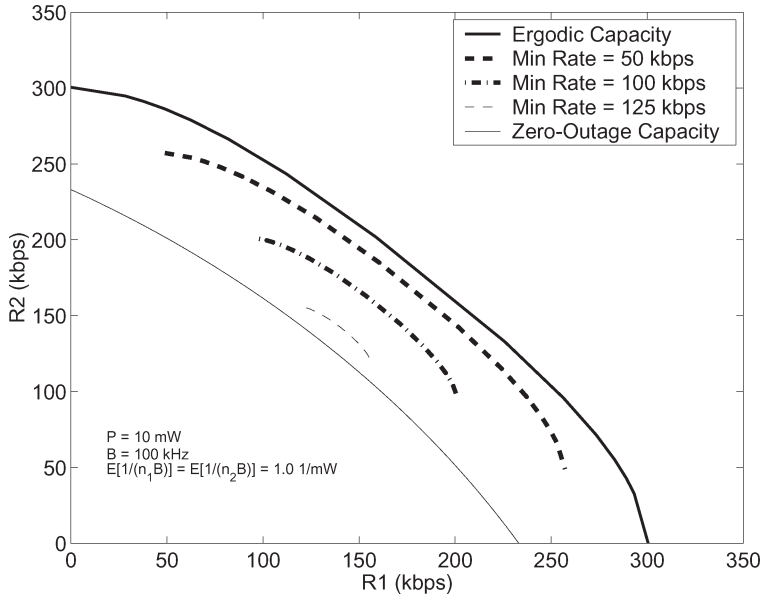


Figure 14.15: Ergodic, zero-outage, and minimum rate BC capacity regions (Rician fading with a K-factor of l, average SNR $= 10$ dB).

The notions of ergodic capacity and outage capacity can also be combined. This combination results in the minimum rate capacity region [50]. A rate vector in this region characterizes the set of all average rate vectors that can be maintained, averaged over all fading states, subject to some minimum rate vector that must be maintained in all states (possible subject to some outage probability). Minimum rate capacity is useful for systems supporting

a mix of delay-constrained and delay-unconstrained data. The minimum rates dictate the data rates available for the constrained data that must be maintained in all fading states, while the rates above these minimums are what is available for the unconstrained data, where these additional rates vary depending on the current fading state. The minimum rate capacity region (with zero outage probability) lies between that of the zero-outage capacity region and the ergodic capacity region: for minimum rates of zero it equals the ergodic capacity region; for minimum rates on the boundary of the zero-outage capacity region, it cannot exceed these boundary points. This is illustrated in Figure 14.15, where we plot the ergodic, zero-outage, and minimum rate capacity region for a BC with Rician fading. We see from this figure that the ergodic capacity region is the largest, since it can adapt to the different channel states in order to maximize its average rate (averaged over all fading states). The zero-outage capacity region is the smallest, since it is forced to maintain a fixed rate in all states, which consumes much power when the fading is severe. The minimum rate capacity region lies between the other two and depends on the minimum rate requirements. As the minimum rate vector that must be maintained in all fading states increases, the minimum rate capacity region approaches the zero-outage capacity region, and as this minimum rate vector decreases, the minimum rate capacity region approaches the ergodic capacity region.

14.6.5 Capacity with Multiple Antennas

We now investigate the capacity region for a BC with multiple antennas. We have seen in Section 10.3 that multiple-input multiple-out (MIMO) systems can provide large capacity increases for single-user systems. The same will be true of multiuser systems: in fact, multiple users can exploit multiple spatial dimensions even more effectively than a single user.

Consider a K-user BC in which the transmitter has M_t antennas and each receiver has M_r antennas. The $M_r \times M_t$ channel matrix $\mathbf{H}_k$ characterizes the channel gains between each antenna at the transmitter and each antenna at the kth receiver. The received signal for the kth user is then

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k, \tag{14.45}$$

where $\mathbf{x}$ is the input to the transmit antennas, and we denote its covariance matrix as $\mathbf{Q}_\mathbf{x} = \mathbf{E}[\mathbf{x}\mathbf{x}^H]$. For simplicity, we normalize the bandwidth to unity,[8] $B = 1$ Hz, and assume the noise vector $\mathbf{n}_k$ is circularly symmetric complex Gaussian with $\mathbf{n}_k \sim N(0, \mathbf{I})$.

When the transmitter has more than one antenna, $M_t > 1$, the BC is no longer degraded. In other words, receivers cannot generally be ranked by their channel quality, since receivers have different channel gains associated with the different antennas at the transmitter. The capacity region of the general non-degraded broadcast channels is unknown. However, an achievable region for this channel was proposed in [52, 53] that was later shown [54] to equal the capacity region. The region is based on the notion of *dirty paper coding* (DPC) [55]. The basic premise of DPC is as follows. If the transmitter (but not the receiver) has perfect, noncausal knowledge of interference to a given user, then the capacity of the channel is the same as if there were no interference or, equivalently, as if the receiver had knowledge of the interference and could subtract it out. Dirty paper coding is a technique that allows noncausally known interference to be "presubtracted" at the transmitter but in such a way that the transmit power is not increased. A more practical (and more general) technique to perform this presubtraction is described in [56].

In the MIMO broadcast channel, DPC can be applied at the transmitter when choosing codewords for different users. The transmitter first picks a codeword $\mathbf{x}_1$ for user 1. The transmitter then chooses a codeword $\mathbf{x}_2$ for user 2 with full (noncausal) knowledge of the codeword intended for user 1. Hence the codeword of user 1 can be presubtracted such that user 2 does not see the codeword intended for user 1 as interference. Similarly, the codeword $\mathbf{x}_3$ for user 3 is chosen such that user 3 does not see the signals intended for users 1 and 2 as interference.

[8] The capacity of unity-bandwidth MIMO channels has a factor of .5 preceding the log function for real (one-dimensional) channels, with no such factor for complex (two-dimensional) channels [51, Chap. 3.1].

This process continues for all K users. The ordering of the users clearly matters in such a procedure and needs to be optimized in the capacity calculation. Let $\pi(\cdot)$ denote a permutation of the user indices and let $\mathbf{Q} = [\mathbf{Q}_1, \ldots, \mathbf{Q}_K]$ denote a set of positive semidefinite covariance matrices with Tr $[\mathbf{Q}_1 + \cdots + \mathbf{Q}_K] \leq P$. Under DPC, if user $\pi(1)$ is encoded first, followed by user $\pi(2)$, and so forth, then the following rate vector is achievable:

$$\mathbf{R}(\pi, \mathbf{Q}) = (R_{\pi(1)}, \ldots, R_{\pi(K)}) : \\ R_{\pi(k)} = \log_2 \frac{\det[\mathbf{I} + \mathbf{H}_{\pi(k)}(\sum_{j \geq k} \mathbf{Q}_{\pi(j)})\mathbf{H}_{\pi(k)}^H]}{\det[\mathbf{I} + \mathbf{H}_{\pi(k)}(\sum_{j > k} \mathbf{Q}_{\pi(j)})\mathbf{H}_{\pi(k)}^H]}, \; k = 1, \ldots, K. \tag{14.46}$$

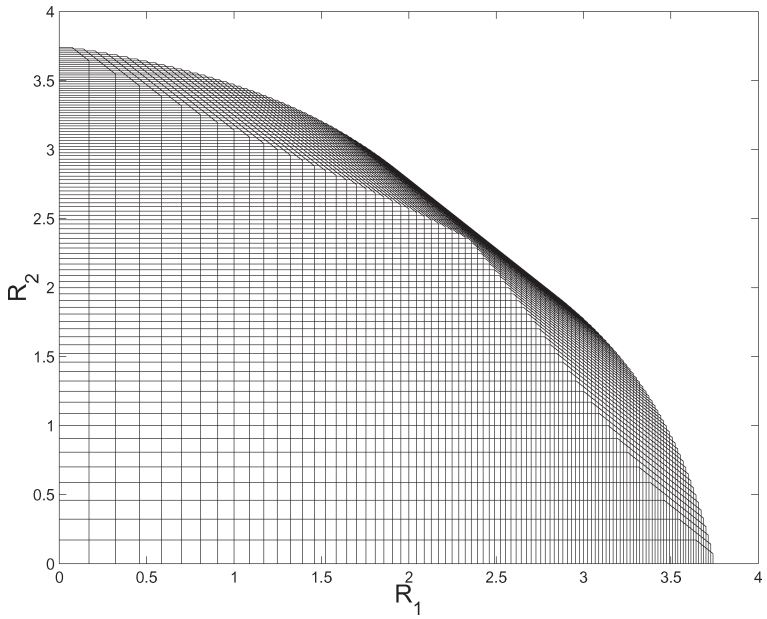


Figure 14.16: MIMO BC capacity region with $\mathbf{H}_1 = [1\ 0.5], \mathbf{H}_2 = [0.5\ 1]$, and $P = 10$.

The capacity region $\mathcal{C}$ is then the convex hull of the union of all such rate vectors over all permutations and all positive semidefinite covariance matrices satisfying the average power constraint:

$$\mathcal{C}_{\mathrm{BC}}(P, \mathbf{H}) \triangleq \mathrm{Co}\left(\bigcup_{\pi, Q} \mathbf{R}(\pi, \mathbf{Q})\right), \tag{14.47}$$

where $\mathbf{R}(\pi, \mathbf{Q})$ is given by (14.46). The transmitted signal is $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_K$ and the input covariance matrices are of the form $\mathbf{Q}_k = \mathbf{E}[\mathbf{x}_k \mathbf{x}_k^H]$. The DPC implies that $\mathbf{x}_1, \ldots, \mathbf{x}_K$ are uncorrelated, and thus Tr $[\mathbf{Q}_\mathbf{x}]$ = Tr $[\mathbf{Q}_1 + \cdots + \mathbf{Q}_K] \leq P$.

One important feature to notice about the rate equations defined by (14.46) is that these equations are neither a concave nor a convex function of the covariance matrices. This makes finding the capacity region difficult, because generally the entire space of covariance matrices that meet the power constraint must be searched [52, 53]. However, as described in Section 14.8, there is a duality between the MIMO BC and the MIMO MAC that can be exploited to greatly simplify this calculation. Figure 14.16 shows the capacity region for a two-user channel computed by exploiting this duality with $M_t = 2$ and $M_r = 1$. The region is defined by the outer boundary, and the lines inside this boundary each corresponds to the capacity region of a different dual MIMO MAC channel

whose sum power equals the power of the MIMO BC. The union of these dual regions yields the boundary of the MIMO BC region, as will be discussed in Section 14.8.

14.7 Uplink (Multiple Access) Channel Capacity

14.7.1 Capacity in AWGN

The MAC consists of K transmitters, each with power P_k, sending to a receiver over a channel with power gain g_k. We assume that all transmitters and the receiver have a single antenna. The received signal is corrupted by AWGN with PSD $N_0/2$. The two-user MAC capacity region is the closed convex hull of all vectors (R_1, R_2) satisfying the following constraints [41, Chap. 14.1]:

$$R_k \leq B \log_2 \left(1 + \frac{g_k P_k}{N_0 B}\right), \; k = 1, 2, \tag{14.48}$$

$$R_1 + R_2 \leq B \log_2 \left(1 + \frac{g_1 P_1 + g_2 P_2}{N_0 B}\right). \tag{14.49}$$

The first constraint (14.48) is just the capacity associated with each individual channel. The second constraint (14.49) indicates that the sum of rates for all users cannot exceed the capacity of a "superuser" with received power equal to the sum of received powers from all users. For K users, the region becomes

$$\mathcal{C}_{\text{MAC}} = \left\{ (R_1, \; \ldots, \; R_K) : \sum_{k \in S} R_k \leq B \log_2 \left(1 + \frac{\sum_{k \in S} g_k P_k}{N_0 B}\right), \; \forall S \subset \{1, 2, \; \ldots, \; K\} \right\}. \tag{14.50}$$

Thus, the region described by (14.50) demonstrates that the sum of rates for any subset of the K users cannot exceed the capacity of a superuser with received power equal to the sum of received powers associated with this user subset.

The sum-rate capacity of a MAC is the maximum sum of rates $\sum_{k=1}^{K} R_k$, where the maximum is taken over all rate vectors $(R_1, \ldots, R_K)$ in the MAC capacity region. As with the like capacity for the BC, the MAC sum-rate also measures the maximum throughput of the system regardless of fairness and is easier to characterize than the K-dimensional capacity region. It can be shown from (14.50) that sum-rate capacity is achieved on the AWGN MAC by having all users transmit at their maximum power, which yields

$$C_{\text{MACSR}} = B \log_2 \left(1 + \frac{\sum_{k=1}^{K} g_k P_k}{N_0 B}\right). \tag{14.51}$$

The intuition behind this result is that each user in the MAC has an individual power constraint, so not allowing a user to transmit at full power wastes system power. By contrast, the AWGN BC sum-rate capacity (14.25) is achieved by transmitting only to the user with the best channel. However, since all users share the power resource, no power is wasted in this case.

The MAC capacity region for two users is shown in Figure 14.17, where C_k and C_k^* are given by

$$C_k = B\log_2\left(1+\frac{g_k P_k}{N_0 B}\right),\ k=1,2; \tag{14.52}$$

$$C_1^* = B\log_2\left(1+\frac{g_1 P_1}{N_0 B + g_2 P_2}\right), \tag{14.53}$$

$$C_2^* = B\log_2\left(1+\frac{g_2 P_2}{N_0 B + g_1 P_1}\right). \tag{14.54}$$

The point $(C_1, 0)$ is the achievable rate vector when transmitter 1 is sending at its maximum rate and transmitter 2 is silent, and the opposite scenario achieves the rate vector $(0, C_2)$. The corner points $(C_1,\ C_2^*)$ and $(C_1^*,\ C_2)$ are achieved using the successive interference cancellation described previously for superposition codes. Specifically, let the first user operate at the maximum data rate C_1. Then its signal will appear as noise to user 2; thus, user 2 can send data at rate C_2^* that can be decoded at the receiver with arbitrarily small error probability. If the receiver then subtracts out user 2's message from its received signal, the remaining message component is just user 1's message corrupted by noise, so rate C_1 can be achieved with arbitrarily small error probability. Hence, $(C_1,\ C_2^*)$ is an achievable rate vector. A similar argument with the user roles reversed yields the rate point $(C_1^*,\ C_2)$. Time sharing between these two strategies yields any point on the straight line connecting $(C_1,\ C_2^*)$ and $(C_1^*,\ C_2)$. This time-sharing line can also be achieved via a technique called *rate splitting*, whereby a given user splits his data stream into multiple substreams and encodes these substreams as if they originated from different virtual users [57]. Note that in the BC the user with the highest SNR must always be decoded last, whereas with MACs the decoding can be done in either order. This is a fundamental difference between the two channels.

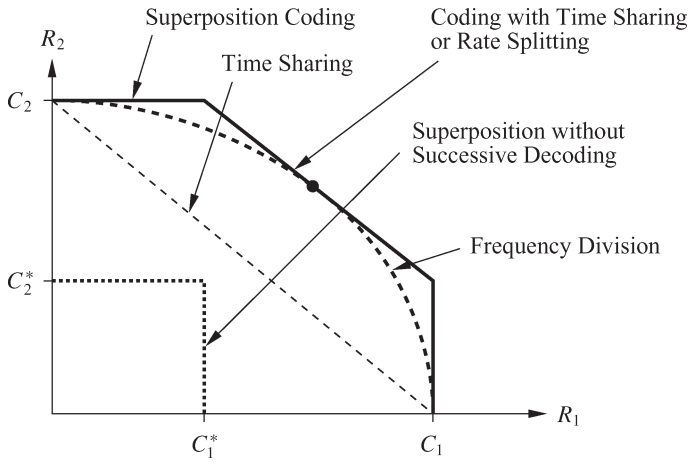


Figure 14.17: Two-user MAC capacity region.

A time-division strategy between two transmitters operating at their maximum rates, given by (14.52), yields any rate vector on the straight line connecting C_1 and C_2. With frequency division, the rates depend on the fraction of the total bandwidth that is allocated to each transmitter. Letting B_1 and B_2 denote the bandwidth allocated to each of the two users, we obtain the achievable rate region

$$\mathcal{C}_{\rm FD} = \bigcup_{\{B_1,B_2:B_1+B_2=B\}} \left(R_1 = B_1\log_2\left(1+\frac{g_1P_1}{N_0B_1}\right), R_2 = B_2\log_2\left(1+\frac{g_2P_2}{N_0B_2}\right)\right), \tag{14.55}$$

which is plotted in Figure 14.17. Clearly this region dominates TD, since setting $B_1 = \tau B$ and $B_2 = (1-\tau)B$ in (14.55) yields $R_1 > \tau C_1$ and $R_2 > (1-\tau)C_2$. It can be shown [41, Chap. 14.3] that this curve touches the capacity

region boundary at one point: this point corresponds to the rate vector that maximizes the sum-rate $R_1 + R_2$. To achieve this rate point, the bandwidths B_1 and B_2 must be proportional to their corresponding received powers $g_1 P_1$ and $g_2 P_2$.

As with the broadcast channel, we can obtain the same achievable rate region with TD as with FD by efficient use of the transmit power. If we take the constraints P_1 and P_2 to be average power constraints then it follows that, since user k uses the channel only a fraction τ_k of the time, its average power over that time fraction can be increased to P_k/τ_k. The rate region achievable through variable-power TD is then given by

$$\mathcal{C}_{\mathrm{TD,VP}} = \bigcup_{\{\tau_1,\tau_2:\tau_1+\tau_2=1\}} \left(R_1 = \tau_1 B \log_2\left(1 + \frac{g_1 P_1}{N_0 \tau_1 B}\right), R_2 = \tau_2 B \log_2\left(1 + \frac{g_2 P_2}{N_0 \tau_2 B}\right)\right), \tag{14.56}$$

and substituting $B_k \triangleq \tau_k B$ in (14.56) yields the same rate region as in (14.55).

Superposition codes without successive decoding can also be used. With this approach, each transmitter's message acts as noise to the others. Thus, the maximum achievable rate in this case cannot exceed $(C_1^*,\ C_2^*)$, which is clearly dominated by FD and TD for some bandwidth or time allocations – in particular, the allocation that intersects the capacity region boundary.

Example 14.8: Consider a MAC channel in AWGN with transmit power $P_1 = P_2 = 100$ mW for both users and with channel gains of $g_1 = .08$ for user 1 and $g_2 = .001$ for user 2. Assume the receiver noise has $N_0 = 10^{-9}$ W/Hz and the system bandwidth is $B = 100$ kHz. Find the corner points of the MAC capacity region. Also find the rate that user 1 can achieve if user 2 requires a rate of $R_2 = 100$ kbps and of $R_2 = 50$ kbps.
Solution: From (14.52)–(14.54) we have:

$$C_1 = B\ \log_2\left(1 + \frac{g_1 P_1}{N_0 B}\right) = 6.34 \cdot 10^5, \tag{14.57}$$

$$C_2 = B\ \log_2\left(1 + \frac{g_2 P_2}{N_0 B}\right) = 1 \cdot 10^5; \tag{14.58}$$

$$C_1^* = B\ \log_2\left(1 + \frac{g_1 P_1}{N_0 B + g_2 P_2}\right) = 5.36 \cdot 10^5, \tag{14.59}$$

$$C_2^* = B\ \log_2\left(1 + \frac{g_2 P_2}{N_0 B + g_1 P_1}\right) = 1.77 \cdot 10^3. \tag{14.60}$$

The maximum rate for user 2 is 100 kbps, so if he requires $R_2 = 100$ kbps then this rate point is associated with the corner point $(C_1^*,\ C_2)$ of the capacity region; hence user 1 can achieve a rate of $R_1 = C_1^* = 536$ kbps. If user 2 requires only $R_2 = 50$ kbps then the rate point lies on the TD portion of the capacity region. In particular, time sharing between the corner points $(C_1,\ C_2^*)$ and $(C_2,\ C_1^*)$ yields, for any τ with $0 \le \tau \le 1$, the rate point $(R_1, R_2) = \tau(C_1, C_2^*) + (1-\tau)(C_1^*,\ C_2)$. The time-share value τ that yields $R_2 = 50$ kbps thus satisfies $\tau C_2^* + (1-\tau)C_2 = R_2$. Solving for τ yields $\tau = (R_2 - C_2)/(C_2^* - C_2) = .51$, about halfway between the two corner points. Then user 1 can achieve rate $R_1 = \tau C_1 + (1-\tau)C_1^* = 5.86 \cdot 10^5$. This example illustrates the dramatic impact of the near–far effect in MAC channels. Even though both users have the same transmit power, the channel gain of user 2 is much less than the gain of user 1. Hence, user 2 can achieve at most a rate of 100 kbps, whereas user 1 can achieve a rate between 536 and 634 kbps. Moreover, the interference from user 2 has little impact on user 1 owing to the weak channel gain associated with the interference: user 1 sees data rates of $C_1 = 634$ kbps without interference and $C_1^* = 536$ kbps with interference. However, the interference from user 1 severely limits the data rate of user 2, decreasing it by almost two orders of magnitude from $C_2 = 100$ kbps to $C_2^* = 1.77$ kbps.

14.7.2 Capacity in Fading

We now consider the capacity region of a multiple access channel with AWGN and fading, where the channel gains for each user change over time. We assume that all transmitters and the receiver have a single antenna and that the receiver has AWGN with PSD $N_0/2$. Each user has an individual power constraint $\bar{P}_k,\ k = 1, \ldots, K$. The time-varying power gain of user k's channel at time i is $g_k[i]$ and is independent of the fading of other users. We define the fading state at time i as $\mathbf{g}[i] = (g_1[i],\ \ldots,\ g_K[i])$, with the time reference dropped when the context is clear. We assume perfect CSI about the fading state at both the transmitter and receiver; the case of receiver CSI only is treated in [58, Chap. 6.3]. Like the broadcast and single-user channels, the fading MAC also has two notions of capacity: the ergodic capacity region, which characterizes the achievable rate vectors averaged over all fading states; and the outage capacity region, which characterizes the maximum rate vector that can be maintained in all states with some nonzero probability of outage.

We first consider the ergodic capacity region as derived in [46]. Define a power policy $\mathcal{P}$ as a function that maps a fading state $\mathbf{g} = (g_1, \ldots, g_K)$ to a set of powers $P_1(\mathbf{g}), \ldots, P_K(\mathbf{g})$, one for each user. Let $\mathcal{F}_{\mathrm{MAC}}$ denote the set of all power policies satisfying the average per-user power constraint $\bar{P}_k$:

$$\mathcal{F}_{\mathrm{MAC}} \stackrel{\triangle}{=} \{\mathcal{P} : \mathbf{E}_{\mathrm{g}}[P_k(\mathbf{g})] \leq \bar{P}_k, k = 1, \ldots, K\}. \tag{14.61}$$

The MAC capacity region assuming a constant fading state $\mathbf{g}$ with power allocation $P_1(\mathbf{g}), \ldots, P_K(\mathbf{g})$ is given by

$$\begin{aligned}&\mathcal{C}_{\mathrm{MAC}}(P_1(\mathbf{g}), \ldots, P_K(\mathbf{g})) \\ &\quad = \left\{(R_1, \ldots, R_K) : \sum_{k\in S} R_k \leq B\ \log_2\left(1 + \frac{\sum_{k\in S} g_k P_k(\mathbf{g})}{N_0 B}\right),\ \forall S \subset \{1, 2,\ \ldots,\ K\}\right\}.\end{aligned} \tag{14.62}$$

The set of achievable rates averaged over all fading states under power policy $\mathcal{P}$ is given by

$$\begin{aligned}&\mathcal{C}_{\mathrm{MAC}}(\mathcal{P}) \\ &\quad = \left\{(R_1,\ \ldots,\ R_K) : \sum_{k\in S} R_k \leq \mathbf{E}_{\mathbf{g}}\left[B\ \log_2\left(1 + \frac{\sum_{k\in S} g_k P_k(\mathbf{g})}{N_0 B}\right)\right],\ \forall S \subset \{1, 2,\ \ldots,\ K\}\right\}.\end{aligned} \tag{14.63}$$

The ergodic capacity region is then the union over all power policies that satisfy the individual user power constraints:

$$\mathcal{C}_{\mathrm{MAC}}(\bar{P}_1,\ \ldots, \bar{P}_K) = \bigcup_{\mathcal{P}\in\mathcal{F}_{\mathrm{MAC}}} \mathcal{C}_{\mathrm{MAC}}(\mathcal{P}). \tag{14.64}$$

From (14.51), (14.63), and (14.64), the sum-rate capacity of the MAC in fading reduces to

$$\mathcal{C}_{\mathrm{MACSR}} = \max_{\mathcal{P}\in\mathcal{F}_{\mathrm{MAC}}} \mathbf{E}_{\mathrm{g}}\left[B \log_2\left(1 + \frac{\sum_{k=1}^{K} g_k P_k(\mathbf{g})}{N_0 B}\right)\right]. \tag{14.65}$$

The maximization in (14.65) is solved using Lagrangian techniques, and the solution reveals that the optimal transmission strategy to achieve sum-rate is to allow only one user to transmit in every fading state [59]. Under this optimal policy, the user that transmits in a given fading state $\mathbf{g}$ is the one with the largest *weighted* channel gain g_k/λ_k, where λ_k is the Lagrange multiplier associated with the average power constraint of the kth user.

This Lagrangian is a function of the user's average power constraint and fading distribution. By symmetry, if all users have the same fading distribution and the same average power constraint, then each user has the same λ_k and so the optimal policy is to allow only the user with the best channel g_k to transmit in fading state $\mathbf{g}$. Once it is determined which user should transmit in a given state, the power the user allocates to that state is determined via a water-filling over time. The intuition behind allowing only one user at a time to transmit is as follows. Since users can adapt their powers over time, system resources are best utilized by assigning them to the user with the best channel and allowing that user to transmit at a power commensurate with his channel quality. But when users have unequal average received power this strategy is no longer optimal: users with weak average received SNR would rarely transmit, so their individual power resources would be underutilized.

The MAC zero-outage capacity region, derived in [47], defines the set of rates that can be simultaneously achieved for all users in *all* fading states while meeting the average power constraints of each user. From (14.50), given a power policy $\mathcal{P}$ that maps fading states to user powers, the MAC capacity region in state $\mathbf{g}$ is

$$\mathcal{C}_{\mathrm{MAC}}(\mathcal{P}) = \left\{ (R_1, \ldots, R_K) : \sum_{k \in S} R_k \leq B \log_2 \left(1 + \frac{\sum_{k \in S} g_k P_k(\mathbf{g})}{N_0 B} \right), \forall S \subset \{1, 2, \ldots, K\} \right\}. \quad (14.66)$$

Then, under policy $\mathcal{P}$, the set of rates that can be maintained in all fading states $\mathbf{g}$ is

$$\mathcal{C}^0_{\mathrm{MAC}}(\mathcal{P}) = \bigcap_{\mathrm{g}} \mathcal{C}_{\mathrm{MAC}}(\mathcal{P}). \quad (14.67)$$

The zero-outage capacity region is then the union of $\mathcal{C}^0_{\mathrm{MAC}}(\mathcal{P})$ over all power policies $\mathcal{P} \in \mathcal{F}_{\mathrm{MAC}}$ that satisfy the user power constraints. Thus, the zero-outage MAC capacity region is given by

$$\mathcal{C}^0_{\mathrm{MAC}}(\bar{P}_1, \ldots, \bar{P}_K) = \bigcup_{\mathcal{P} \in \mathcal{F}_{\mathrm{MAC}}} \bigcap_{\mathrm{g}} \mathcal{C}_{\mathrm{MAC}}(\mathcal{P}). \quad (14.68)$$

The outage capacity region of the MAC is similar to the zero-outage capacity region, except that users can suspend transmission in some outage states subject to a given nonzero probability of outage. As with the BC, the MAC outage capacity region is more difficult to obtain than the zero-outage capacity region, since in any given fading state the transmission strategy must determine which users to put into outage, the decoding order of the non-outage users, and the power at which these non-outage users should transmit. The MAC outage capacity region is obtained implicitly in [49] by determining whether a given rate vector $\mathbf{R}$ can be maintained in all fading states – subject to a given per-user outage probability – without violating the per-user power constraints. Ergodic and outage capacities can also be combined to obtain the minimum rate capacity region for the MAC. As with the BC, this region characterizes the set of all average rate vectors (averaged over all fading states) that can be maintained subject to some minimum rate vector that must be maintained in all states with some outage probability (possibly zero). The minimum rate capacity region for the fading MAC is derived in [60] using the duality principle that relates capacity regions of the BC and the MAC. This duality principle is described in the next section.

14.7.3 Capacity with Multiple Antennas

We now consider MAC channels with multiple antennas. As in the MIMO BC model, we normalize bandwidth to unity, $B = 1$ Hz, and assume the noise vector $\mathbf{n}$ at the MAC receiver is circularly symmetric complex Gaussian with $\mathbf{n} \sim N(0, \mathbf{I})$. Let the $M_r \times M_t$ matrix $\mathbf{H}_k$ characterize the MIMO MAC channel gains between each antenna at the kth user's transmitter and each antenna at the receiver. Define $\mathbf{H} = [\mathbf{H}_1 \ldots \mathbf{H}_K]$. Then the capacity region of the Gaussian MIMO MAC where user k has channel gain matrix $\mathbf{H}_k$ and power P_k is given [61, 62, 63] by

$$\mathcal{C}_{\mathrm{MAC}}((P_1 \ldots, P_K); \mathbf{H})$$
$$= \bigcup_{\{\mathbf{Q}_k \geq 0, \mathrm{Tr}(\mathbf{Q}_k) \leq P_k,\ \forall k\}} \Bigg\{ (R_1,\ \ldots,\ R_K) : \sum_{k \in S} R_k \leq \log_2 \det \left[\mathbf{I} + \sum_{k \in S} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H \right],\ \forall S \subseteq \{1, \ldots, K\} \Bigg\}. \quad (14.69)$$

This region is achieved as follows. The kth user transmits a zero-mean Gaussian random vector $\mathbf{x}_k$ with spatial covariance matrix $\mathbf{Q}_k = \mathbf{E}[\mathbf{x}_k \mathbf{x}_k^H]$. Each set of covariance matrices $(\mathbf{Q}_1, \ldots, \mathbf{Q}_K)$ corresponds to a K-dimensional polyhedron (i.e., $\{(R_1,\ \ldots,\ R_K) : \sum_{k \in S} R_k \leq \log_2 \det[\mathbf{I} + \sum_{k \in S} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H],\ \forall S \subseteq \{1, \ldots, K\}\}$), and the capacity region is equal to the union (over all covariance matrices satisfying the power constraints) of all such polyhedrons. The corner points of this pentagon can be achieved by successive decoding: users' signals are successively decoded and subtracted out of the received signal. Note that the capacity region (14.69) has several similarities with its single-antenna counterpart: it is defined based on the rate sum associated with subsets of users, and the corner points of the region are obtained using successive decoding.

For the two-user case, each set of covariance matrices corresponds to a pentagon, which is similar in form to the capacity region of the single-antenna MAC. For example, the corner point where $R_1 = \log_2 \det[\mathbf{I} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^H]$ and

$$\begin{aligned} R_2 &= \log_2 \det[\mathbf{I} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^H + \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^H] - R_1 \\ &= \log_2 \det[\mathbf{I} + (\mathbf{I} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^H)^{-1} \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^H] \end{aligned} \quad (14.70)$$

corresponds to decoding user 2 first (i.e., in the presence of interference from user 1) and decoding user 1 last (without interference from user 2).

14.8 Uplink-Downlink Duality

The downlink and uplink channels shown in Figure 14.1 appear to be quite similar: the downlink is almost the same as the uplink with the direction of the arrows reversed. However, there are three fundamental differences between the two channel models. First, in the downlink there is an additive noise term associated with each receiver, whereas in the uplink there is only one additive noise term (because there is only one receiver). A second fundamental difference is that the downlink has a single power constraint associated with the transmitter, whereas the uplink has different power constraints associated with each user. Finally, on the downlink both the signal and interference associated with each user travel through the same channel, whereas on the uplink these signals travel through different channels, which gives rise to the near–far effect. Despite extensive study of uplink and downlink channels individually, there has been little effort to draw connections between the two models or exploit these connections in analysis and design. In this section we describe a duality relationship between the two channels, and we show how this relationship can be used in capacity analysis and in the design of uplink and downlink transmission strategies.

We say that K-user downlink and uplink, as shown in Figure 14.1 for $K = 3$, are *duals* of each other if all of the following three conditions hold.

1. The channel impulse responses $h_k(t)$, $k = 1, \ldots, K$, in the downlink are the same as in the uplink for all k. For AWGN or flat fading channels, this implies that bi-directional links have the same channel gain in each direction.

2. Each channel in the downlink has the same noise statistics, and these statistics are the same as those of the channel noise in the uplink.

3. The power constraint P on the downlink equals the sum of individual power constraints P_k, $k = 1, \ldots, K$, on the uplink.

Despite the similarities between the downlink (BC) and uplink (MAC), their capacity regions are quite different. In particular, the two-user AWGN BC capacity region (shown by the largest rate region in Figure 14.13 is markedly different from the two-user AWGN MAC capacity region (shown in Figure 14.17). The capacity regions of dual MACs and BCs are also very different in fading under any of the fading channel capacity definitions: ergodic, outage, or minimum rate capacity. However, despite their different shapes, the capacity regions of the dual channels are both achieved using a superposition coding strategy, and the optimal decoders for the dual channels exploit successive decoding and interference cancellation.

The duality relationship between the two channels is based on exploiting their similar encoding and decoding strategies while bridging their differences by (i) summing the individual MAC power constraints to obtain the BC power constraint and (ii) scaling the BC gains to achieve the near–far effect of the MAC. This relationship was developed in [60], where it was used to show that the capacity region and optimal transmission strategy of either the BC or the MAC can be obtained from the capacity region and optimal transmission strategy of the dual channel. In particular, it was shown in [60] that the capacity region of the AWGN BC with power P and channel gains $\mathbf{g} = (g_1, \ldots, g_K)$ is equal to the capacity region of the dual AWGN MAC with the same channel gains, but where the MAC is subject to a sum power constraint $\sum_{k=1}^{K} P_k \leq P$ instead of individual power constraints $(P_1, \ldots, P_k)$. The sum power constraint in the MAC implies that the MAC transmitters draw power from a single pooled power source with total power P and that power is allocated between the MAC transmitters such that $\sum_{k=1}^{K} P_k \leq P$. Mathematically, the BC capacity region can be expressed as the union of capacity regions for its dual MAC with a pooled power constraint:

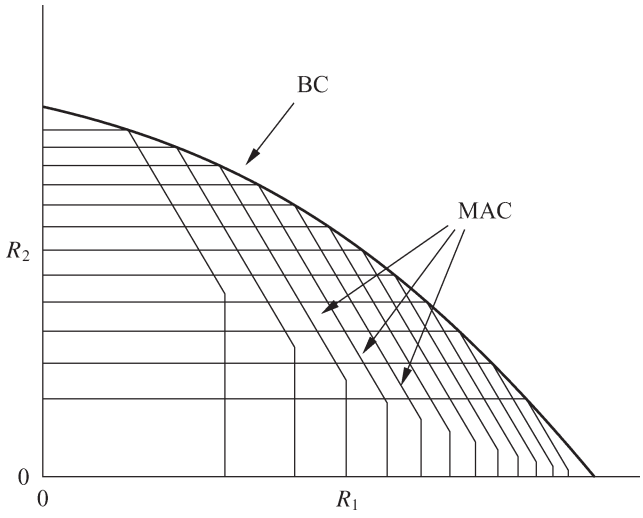


Figure 14.18: AWGN downlink (BC) capacity region as a union of capacity regions for the dual uplink (MAC).

$$\mathcal{C}_{\mathrm{BC}}(P, \mathbf{g}) = \bigcup_{\{(P_1,\ldots,P_K):\Sigma_{i=1}^{K} P_k = P\}} \mathcal{C}_{\mathrm{MAC}}(P_1, \ldots, P_K; \mathbf{g}) \tag{14.71}$$

[60]. Here $\mathcal{C}_{\mathrm{BC}}(P, \mathbf{g})$ is the AWGN BC capacity region with total power constraint P and channel gains $\mathbf{g} = (g_1, \ldots, g_K)$, as given by (14.23) with $n_k = N_0/g_k$, and $\mathcal{C}_{\mathrm{MAC}}(P_1, \ldots, P_K; \mathbf{g})$ is the AWGN MAC capacity

region with individual power constraints $P_1, \ldots, P_K$ and channel gains $\mathbf{g} = (g_1, \ldots, g_K)$, as given by (14.50). This relationship is illustrated for two users in Figure 14.18, where we see the BC capacity region formed from the union of MAC capacity regions with different power allocations between MAC transmitters that sum to the total power P of the dual BC.

In addition to the capacity region relationship of (14.71), it is also shown in [60] that the optimal power allocation for the BC associated with any point on the boundary of its capacity region can be obtained from the allocation of the sum power on the dual MAC that intersects with that point. Moreover, the decoding order of the BC for that intersection point is the reverse decoding order of this dual MAC. Thus, the optimal encoding and decoding strategy for the BC can be obtained from the optimal strategies associated with its dual MAC. This connection between optimal uplink and downlink strategies may have interesting implications for practical designs.

Duality also implies that the MAC capacity region can be obtained from that of its dual BC. This relationship is based on the notion of channel scaling. It is easily seen from (14.50) that the AWGN MAC capacity region is not affected if the kth user's channel gain g_k is scaled by power gain α – as long as its power P_k is also scaled by $1/\alpha$. However, the dual BC is fundamentally changed by channel scaling because the encoding and decoding order of superposition coding on the BC is determined by the order of the channel gains. Thus, the capacity region of the BC with different channel scalings will be different, and it is shown in [60] that the MAC capacity region can be obtained by taking an intersection of the BC with all possible channel scalings α_k on the kth user's channel. Mathematically, we obtain the MAC capacity region from the dual BC as

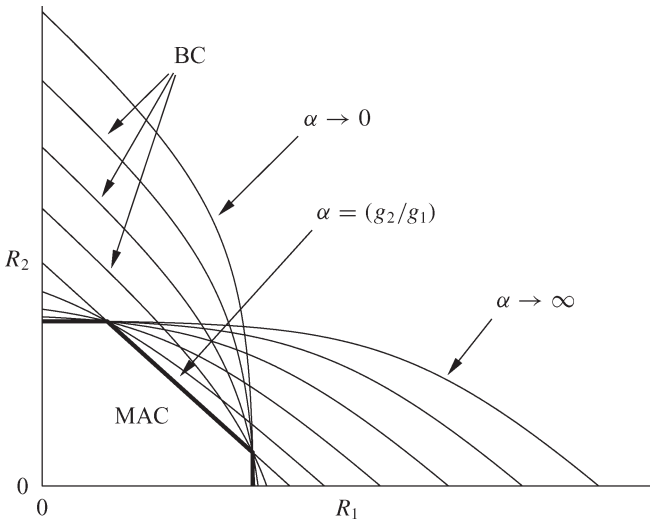


Figure 14.19: AWGN uplink (MAC) capacity region as an intersection of capacity regions for the scaled dual downlink (BC).

$$\mathcal{C}_{\mathrm{MAC}}(P_1, \ldots, P_K; \mathbf{g}) = \bigcap_{(\alpha,\ldots,\alpha_K)>0} \mathcal{C}_{BC}\left(P = \sum_{k=1}^{K} \frac{P_k}{\alpha_k};\ \mathbf{g} = (\alpha_1 g_1, \ldots, \alpha_K g_K)\right), \tag{14.72}$$

where $\mathcal{C}_{\mathrm{MAC}}(P_1,\ \ldots,\ P_K; \mathbf{g})$ and $\mathcal{C}_{\mathrm{BC}}(P, \mathbf{g})$ are defined as in (14.71). This relationship is illustrated in Figure 14.19 for two users with channel gain $\mathbf{g} = (g_1,\ g_2)$. The figure shows that the MAC capacity region is formed from the intersection of BC capacity regions with different channel scalings α applied to the first user.[9] As $\alpha \to 0$, the channel gain $\alpha\, g_1$ of the first user goes to zero but the total power $P = P_1/\alpha + P_2$ goes to infinity. Since user

[9]It is sufficient to take the intersection for scaling over just $K - 1$ users, because scaling by $(\alpha_1, \ldots, \alpha_{K-1},\ \alpha_K)$ is equivalent to scaling by $(\alpha_1/\alpha_K, \ldots, \alpha_{K-1}/\alpha_K, 1)$.

2's channel gain doesn't change, he takes advantage of the increased power and his rate grows asymptotically large with α. The opposite happens as $\alpha \to \infty$: user 1's channel gain grows and the total power $P = P_1/\alpha + P_2 \geq P_2$, so user 1 takes advantage of his increasing channel gain to achieve an asymptotically large rate with any portion of the total power P. All scalings between zero and infinity sketch out different BC capacity regions that intersect to form the MAC region. In particular, when $\alpha = g_2/g_1$, the channel gains of both users in the scaled BC are the same, and this yields the time-sharing segment of the MAC capacity region. The optimal decoding order of the MAC for a given point on its capacity region can also be obtained from the channel scaling associated with the dual scaled BC whose capacity region intersects the MAC capacity region at that point.

These duality relationships are extended in [60] to many other important channel models. In particular, duality applies to flat fading MACs and BCs, so that the ergodic, outage, and minimum rate capacity regions (along with the optimal encoding and decoding strategies) for one channel can be obtained from the regions and strategies for the dual channel. MAC and BC duality also holds for parallel and frequency-selective fading channels, which defines the connection between the capacity regions of frequency-selective fading MACs and BCs [64, 65]. Another important application of duality is to multiple antenna (MIMO) MACs and BCs [54, 66, 67]. In [66] the notion of duality between the BC and the MAC was extended to MIMO systems such that the MIMO BC capacity region with power constraint P and channel gain matrix $\mathbf{H}$ was shown to equal the union of capacity regions of the dual MAC with channel gain matrix $\mathbf{H}^H$, where the union is taken over all individual power constraints that sum to P. Mathematically, we have

$$\mathcal{C}_{\mathrm{BC}}(P, \mathbf{H}) = \bigcup_{(P_1,\ldots,P_K):\Sigma_{k=1}^{K} P_k = P} \mathcal{C}_{\mathrm{MAC}}((P_1, \ldots, P_K); \mathbf{H}^H), \tag{14.73}$$

where $\mathcal{C}_{\mathrm{BC}}(P, \mathbf{H})$ is given by (14.47) and $\mathcal{C}_{\mathrm{MAC}}((P_1, \ldots, P_K); \mathbf{H}^H)$ is given by (14.69). This duality relationship is illustrated in Figure 14.16, where the MIMO BC capacity region is defined by the outer boundary in the figure. The regions inside this boundary are the MIMO MAC capacity region under different individual user power constraints that sum to the total BC power P. Recall that the MIMO BC capacity region is extremely difficult to compute directly, since it is neither concave nor convex over the covariance matrices that must be optimized. In contrast, the optimal MIMO MAC is obtained via a standard convex optimization that is easy to solve [68]. Moreover, duality not only relates the two capacity regions but can also be used to obtain the optimal transmission strategy for a point in the MIMO BC capacity region from a duality transformation of the optimal MIMO MAC strategy that achieves the same point. Thus, for MIMO channels, duality can be exploited to simplify calculation of the capacity region and also to simplify finding the corresponding optimal transmission strategy.

14.9 Multiuser Diversity

Multiuser diversity takes advantage of the fact that, in a system with many users whose channels fade independently, at any given time some users will have better channels than others. By transmitting only to users with the best channels at any given time, system resources are allocated to the users that can best exploit them, which leads to improved system capacity and/or performance. Multiuser diversity was first explored in [59] as a means to increase throughput and reduce error probability in uplink channels, and the same ideas can be applied to downlink channels. The multiuser diversity concept is an extension of the single-user diversity concepts described in Chapter 7. In single-user diversity systems, a point-to-point link consists of multiple independent channels whose signals can be combined to improve performance. In multiuser diversity the multiple channels are associated with different users, and the system typically uses selection diversity to select the user with the best channel in any given fading state. The multiuser diversity gain relies on disparate channels between users and so, the larger the dynamic range of the fading, the higher the multiuser diversity gain. In addition, as with any diversity technique, performance

improves with the number of independent channels. Thus, multiuser diversity is most effective in systems with a large number of users.

From Section 14.6, we have seen that the total throughput (sum-rate capacity) of the fading downlink is maximized by allocating the full system bandwidth to the user with the best channel in each fading state. As described in Section 14.7, a similar result holds for the fading uplink if all users have the same fading distribution and average power. If the users have different fading statistics or average powers, then the channel in any given state is allocated to the user with the best weighted channel gain, where the weight depends on the user's channel gain in the given state, his fading statistics, and his average power constraint. The notion of scheduling transmissions to users based on their channel conditions is called *opportunistic scheduling*, and numerical results in [40, 59] show that opportunistic scheduling, when coupled with power control, can significantly increase both uplink and downlink throughput as measured by sum-rate capacity.

Opportunistic scheduling can also improve BER performance [59]. Let $\gamma_k[i]$ for $k = 1, \ldots, K$ denote the SNR for each user's channel at time i. By transmitting only to the user with the largest SNR, the system SNR at time i is $\gamma[i] = \max_k \gamma_k[i]$. It is shown in [69] that, in independent and identically distributed Rayleigh fading, this maximum SNR is roughly $\ln K$ larger than the SNR of any one user as K grows asymptotically large, leading to a multi-user diversity gain in SNR of $\ln K$. Moreover, if $P_s(\gamma)$ denotes the probability of symbol error for the user with the best channel gain at time i, then $P_s(\gamma)$ will exhibit the same diversity gains as selection combining in a single-user system (described in Section 7.2.2) as compared to the probability of error associated with any one user. As the number of users in the system increases, the probability of error approaches that of an AWGN channel without fading, analogous to increasing the number of branches in single-user selection combining diversity.

Scheduling transmission to users with the best channel raises two problems in wireless systems: fairness and delay. If user fade levels change very slowly, then one user will occupy the system for a long period of time. The time between channel uses for any one user could be quite long, and such latency might be unacceptable for a given application. In addition, users with poor average SNRs will rarely have the best channel and thus are rarely allowed to transmit, which leads to unfairness in the allocation of the system resources. A solution to the fairness and delay problems in the downlink, called *proportional fair scheduling*, was proposed in [69]. Suppose that, at time i, each of the K users in the downlink system can support rate $R_k[i]$ if allocated the full power and system bandwidth. Let $T_k[i]$ denote the average throughput of the kth user at time i, averaged using an exponentially weighted low pass filter parameterized by i_c, which is a parameter of the scheduler design. In the ith time slot, the scheduler then transmits to the user with the largest ratio $R_k[i]/T_k[i]$. If at time i all users have experienced the same average throughput $T_k[i] = T[i]$, then this scheduler transmits to the user with the best channel. Suppose, however, that one user (user j) has experienced poor throughput, so that $T_j[i] \ll T_k[i]$ for $j \neq k$. Then, at time i, user j will likely have a high ratio of $R_j[i]/T_j[i]$ and thus will be favored in the allocation of resources at time i. Assuming that at time i the user k^* has the highest ratio of $R_k[i]/T_k[i]$, the throughput on the next times lot is updated as

$$T_k[i+1] = \begin{cases} (1 - 1/i_c)T_k[i] + (1/i_c)R_k[i] & k = k^*, \\ (1 - 1/i_c)T_k[i] & k \neq k^*. \end{cases} \tag{14.74}$$

With this scheduling scheme, users with the best channels are still allocated the channel resources when throughput between users is reasonably fair. However, if the throughput of any one user is poor, that user will be favored for resource allocation until his throughput becomes reasonably balanced with that of the other users. Clearly this scheme will have a lower throughput than allocating all resources to the user with the best channel (which maximizes throughput), and the throughput penalty will increase as the users have more disparate average channel qualities. The latency with this scheduling scheme is controlled via the weighting parameter i_c. As this parameter increases the latency also increases, but system throughput increases as well because the scheduler has more flexibility in allocating resources to users. As i_c grows to infinity, the proportional fair scheduler simply reduces to

allocating system resources to the user with the best channel, assuming all users have the same average SNR. The proportional fair scheduling algorithm is part of the standard for packet data transmission in CDMA2000 cellular systems [70]; its performance for that system is evaluated in [71]. Alternative methods for incorporating fairness and delay constraints in opportunistic scheduling have been evaluated in [72, 73], along with their performance under practical constraints such as imperfect channel estimates.

14.10 Multiuser MIMO Systems

Multiuser systems with multiple antennas at the transmitter(s) and/or receiver(s) are called *Multiuser MIMO systems*. These multiple antennas can significantly enhance performance in many ways. The antennas can be used to provide diversity gain to improve BER performance. The capacity region of the multiuser channel is increased with multiple transmit and receive antennas, providing multiplexing gain. Finally, multiple antennas can provide directivity gain to spatially separate users, which reduces interference. There typically are trade-offs among these three types of performance gains in MIMO multiuser systems [74].

The multiplexing gain of a MIMO multiuser system characterizes the increase in the uplink or downlink capacity region associated with adding multiple antennas. The capacity regions of MIMO multiuser channels have been extensively studied, motivated by the large capacity gains associated with single-user systems. For AWGN channels the MIMO capacity region for both the uplink and the downlink is described (respectively) in Section 14.7.3 and Section 14.6.5. These results can be extended to find the MIMO capacity region in fading with perfect CSI at all transmitters and receivers. Capacity results and open problems related to MIMO multiuser fading channels under other assumptions about channel CSI are described in [75].

Beamforming was discussed in Section 10.4 as a technique to achieve full diversity in single-user systems at the expense of some capacity loss. In multiuser systems, beamforming has less of a capacity penalty because of the multiuser diversity effect, and in fact beamforming can achieve the sum-rate capacity of the MIMO downlink in the asymptotic limit of a large number of users [76, 77].

Multiuser diversity is based on the idea that in multiuser channels the channel quality varies across users, so performance can be improved by allocating system resources at any given time to the users with the best channels. Design techniques to exploit multiuser diversity were discussed in Section 14.9 for single-antenna multiuser systems. In MIMO multiuser systems, the benefits of multiuser diversity are twofold. First, MIMO multiuser diversity provides improved channel quality, since only users with the best channels are allocated system resources. In addition, MIMO multiuser diversity provides abundant directions where users have good channel gains, so that users chosen for resource allocation in a given state have not only good channel quality but also good spatial separation, thereby limiting interference between them. This twofold diversity benefit allows relatively simple, suboptimal transmitter and receiver techniques to achieve near-optimal performance as the number of users increases [76, 78]. It also eliminates the requirement for multiple receive antennas in downlinks and multiple transmit antennas in uplinks in order to obtain large capacity gains, which reduces complexity and power consumption in the mobile terminal. In particular, the sum-rate capacity gain in MIMO BCs increases roughly linearly with the minimum of the number of transmit antennas and the total number of receive antennas associated with all users [79]. Thus, given a fixed number of transmit antennas at a base station and a large number of users, there is no capacity benefit to having multiple antennas at the user terminals. Similarly, the sum-rate capacity gain in MIMO MACs increases roughly linearly with the minimum of the number of receive antennas and the total number of transmit antennas associated with all users. Hence, in a system with many users and a large number of receive antennas at the base station, adding transmit antennas at the user terminals will not increase capacity. Note that multiuser diversity increases with the dynamic range and rate of the channel fading. By modulating (in a controlled fashion) the amplitude and phase of multiple transmit antennas, the fading rate and dynamic range can be increased, leading to higher multiuser diversity gains. This technique, called *opportunistic beamforming*, is investigated in [69].

Space-time modulation and coding techniques for MIMO multiuser systems have also been developed [80, 81, 82]. The goal of these techniques is to achieve the full range of diversity, multiplexing, and directivity trade-offs inherent in MIMO multiuser systems. Multiuser detection techniques can also be extended to MIMO channels and provide substantial performance gains [83, 84, 85]. In wideband channels, the multiuser MIMO techniques must also cope with frequency-selective fading [82, 86, 87]. Advanced transmission techniques for these wideband channels promise even more significant performance gains than in narrowband channels, since frequency-selective fading provides yet another form of diversity. At low SNRs, however, multiple antennas are mainly used to collect energy. The challenge for MIMO multiuser systems is to develop signaling techniques of reasonable complexity that deliver on the promised performance gains even in practical operating environments.

Chapter 14 Problems

1. Consider an FDMA system for multimedia data users. The modulation format requires 10 MHz of spectrum for each user, and guard bands of 1 MHz are required on each side of the allocated spectrum in order to minimize out-of-band interference. What total bandwidth is required to support 100 simultaneous users in this system?

2. Propagation conditions that introduce delay spread and Doppler must separate out time or frequency channels with guard bands to maintain channel orthogonality after signal transmission. For a channel with a 20 Hz Doppler spread and 3 msec delay spread, find the guard bands needed in time for a TDMA systems, and those needed in frequency for an FDMA system, such that the channels remain orthogonal. For a fixed bandwidth of 20 MHz, which system can accommodate a larger number of orthogonal users?

3. Consider a DS CDMA system occupying 10 MHz of spectrum. Assume an interference-limited system with a spreading gain of $G = 100$ and code cross-correlation of $1/G$.

   (a) For the MAC, find a formula for the SIR of the received signal as a function of G and the number of users K. Assume that all users transmit at the same power and that there is perfect power control, so all users have the same received power.

   (b) Based on your SIR formula in part (a), find the maximum number of users K that can be supported in the system, assuming BPSK modulation with a target BER of 10^{-3}. In your BER calculation you can treat interference as AWGN. How does this compare with the maximum number of users K that an FDMA system with the same total bandwidth and information signal bandwidth could support?

   (c) Modify your SIR formula in part (a) to include the effect of voice activity, defined as the percentage of time that users are talking, so interference is multiplied by this percentage. Also find the voice activity factor such that the CDMA system accommodates the same number of users as an FDMA system. Is this a reasonable value for voice activity?

4. Consider an FH CDMA system that uses FSK modulation and the same spreading and information bandwidth as the DS CDMA system in Problem 14.3. Thus, there are $G = 100$ frequency slots in the system, each of bandwidth 100 kHz. The hopping codes are random and uniformly distributed, so the probability that a given user occupies a given frequency slot on any hop is .01. As in Problem 14.3, noise is essentially negligible, so the probability of error on a particular hop is zero if only one user occupies that hop. Also assume perfect power control, so the received power from all users is the same.

   (a) Find an expression for the probability of bit error when m users occupy the same frequency slot.

   (b) Assume there is a total of K users in the system at any time. What is the probability that, on any hop, there is more than one user occupying the same frequency?

   (c) Find an expression for the average probability of bit error as a function of K, the total number of users in the system.

   Consider a multiuser MIMO system downlink where the transmitter has 4 antennas and each user has 4 antennas. For the MIMO channel below associated with user k, how many spatial channels are available to

this user?

$$\mathbf{H} = \begin{bmatrix} .1 & .3 & .7 \\ .2 & & \\ .5 & .4 & .1 \\ .1 & & \\ .2 & .6 & .8 \\ .5 & & \end{bmatrix}. \tag{14.75}$$

5. Consider a single-antenna full duplex system where the TX self-interference is 90 dB larger than the received signal.

   (a) Assume each bit of an A/D converter contributes 6 dB of dynamic range. If the combination of analog isolation and interference cancellation can remove 50 dB of the TX self-interference, what is the smallest A/D converter that preserves the received signal after this analog interference mitigation?

   (b) How much additional digital interference cancellation is needed to achieve a 10^-3 BER for a BPSK modulated signal?

6. Compute the maximum throughput T for a pure ALOHA and a slotted ALOHA random access system, along with the load L that achieves the maximum in each case.

7. Consider a pure ALOHA system with a transmission rate of $R = 10$ Mbps. Compute the load L and throughput T for the system assuming 1000 bit packets and a Poisson arrival rate of $\lambda = 10^3$ packets per second. Also compute the effective data rate (rate of bits successfully received). What other value of load L results in the exact same throughput?

8. Consider a three-user uplink channel with channel power gains $g_1 = 1,\ g_2 = 3$, and $g_3 = 5$ from user k to the receiver, $k = 1, 2, 3$. Assume that all three users require a 10-dB SINR. The receiver noise is $n = 1$.

   (a) Confirm that the vector equation $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ given by (14.11) is equivalent to the SINR constraints of each user.

   (b) Determine whether a feasible power vector exists for this system such that all users meet the required SINR constraints and, if so, find the optimal power vector $\mathbf{P}^*$ such that the desired SINRs are achieved with minimum transmit power. Assume $\rho = .1$.

9. Find the two-user broadcast channel capacity region under superposition coding for transmit power $P = 10\ \mathrm{mW},\ B = 100\ \mathrm{kHz}$, and $N_0 = 10^{-9}\ \mathrm{W/Hz}$.

10. Show that the sum-rate capacity of the AWGN BC is achieved by sending all power to the user with the highest channel gain.

11. Derive a formula for the optimal power allocation on a fading broadcast channel to maximize sum-rate.

12. Find the sum-rate capacity of a two-user fading BC, where the fading on each user's channel is independent. Assume that each user has (i) a received power of 10 mW and (ii) an effective noise power of 1 mW with probability .5 or 5 mW with probability .5.

13. Find the sum-rate capacity for a two-user broadcast fading channel where each user experiences Rayleigh fading. Assume an average received power of $P = 10$ mW for each user, bandwidth $B = 100$ kHz, and $N_0 = 10^{-9}$ W/Hz.

14. Consider the set of achievable rates for a broadcast fading channel under frequency division. Given any rate vector in $C_{\mathrm{FD}}(\mathcal{P},\ B)$ for a given power policy $\mathcal{P}$ and bandwidth allocation policy $\mathcal{B}$, as defined in (14.38), find the timeslot and power allocation policy that achieves the same rate vector.

15. Consider a time-varying broadcast channel with total bandwidth $B = 100$ kHz. The effective noise for user 1 has $n_1 = 10^{-5}$ W/Hz with probability 3/4 or $n_1 = 2 \cdot 10^{-5}$ W/Hz with probability 1/4. The effective noise for user 2 takes the value $n_2 = 10^{-5}$ W/Hz with probability 1/2 or the value $n_2 = 2 \cdot 10^{-5}$ W/Hz with probability 1/2. These noise densities are independent of each other over time. The total transmit power is $P = 10$ W.

    (a) What is the set of all possible joint noise densities and corresponding probabilities?

    (b) Obtain the optimal power allocation between the two users and the corresponding time-varying capacity rate region using time division. Assume that (i) user k is allocated a fixed timeslot τ_k for all time (where $\tau_1 + \tau_2 = 1$) and a fixed average power P over all time and (ii) that each user may change its power within its own timeslot, subject to the average constraint P. Find a rate point that exceeds this region, assuming you don't divide power equally.

    (c) Assume now fixed frequency division, where the bandwidth assigned to each user is fixed and evenly divided between the two users: $B_1 = B_2 = B/2$. Assume also that you allocate half the power to each user within his respective bandwidth ($P_1 = P_2 = P/2$) and that you can vary the power over time, subject only to the average power constraint $P/2$. What is the best rate point that can be achieved? Find a rate point that exceeds this region, assuming that you don't share power and/or bandwidth equally.

    (d) Is the rate point $(R_1 = 100000,\ R_2 = 100000)$ in the zero-outage capacity region of this channel?

16. Show that the K-user AWGN MAC capacity region is not affected if the kth user's channel power gain g_k is scaled by α as long as the kth user's transmit power P_k is also scaled by $1/\alpha$.

17. Consider a multiple access channel being shared by two users. The total system bandwidth is $B = 100$ kHz. The transmit power of user 1 is $P_1 = 3$ mW, while the transmit power of user 2 is $P_2 = 1$ mW. The receiver noise density is .00l μW/ Hz. You can neglect any path loss, fading, or shadowing effects.

    (a) Suppose user 1 requires a data rate of 300 kbps to see videos. What is the maximum rate that can be assigned to user 2 under time division? Under superposition coding with successive interference cancellation?

    (b) Compute the rate pair $(R_1,\ R_2)$ where the frequency-division rate region intersects the region achieved by code division with successive interference cancellation (assuming G = 1).

    (c) Compute the rate pair $(R_1,\ R_2)$ such that $R_1 = R_2$ (i.e., where the two users get the same rate) for time division and for spread-spectrum code division with and without successive interference cancellation for a spreading gain $G = 10$. *Note*: To obtain this region for $G > 1$ you must use the same reasoning on the MAC as was used to obtain the BC capacity region with $G > 1$.

18. Show that the sum-rate capacity of the AWGN MAC is achieved by having all users transmit at full power.

19. Derive the optimal power adaptation for a two-user fading MAC that achieves the sum-rate point.

20. Find the sum-rate capacity of a two-user fading MAC where the fading on each user's channel is independent. Assume each user has (i) a received power of 10 mW and (ii) an effective noise power of 1 mW with probability .5 or 5 mW with probability .5.

21. Consider a three-user fading downlink with bandwidth 100 kHz. Suppose the three users all have the same fading statistics, so that their received SNRs when allocated the full power and bandwidth are 5 dB with probability 1/3, 10 dB with probability 1/3, and 20 dB with probability 1/3. Assume a discrete-time system with i.i.d. fading at each time slot.

    (a) Find the maximum throughput of this system if (at each time instant) the full power and bandwidth are allocated to the user with the best channel.

    (b) **Degradation of WiFi** In a coffee shop, there are M users communicating with a WiFi access point. At every time slot, a user can communicate over any one of N different frequency slots.

        i. If the access point performs centralized control, it can allocate frequency slots to users at each point in time. What is the expected number of users that are served at any time instant?

        ii. WiFi uses decentralized random access protocols. Here, each user chooses a frequency slot at random. If a frequency slot is picked by a single user, it can be used for communication. Otherwise, there are collisions. Find the expected number of users that are served at any time instant. Evaluate this expression for $N = 10$ and $M = 3, 10$ and 30.
        Hint: Use linearity of expectations

        iii. What fraction of users are served when $M = kN$ for $k \geq 1$ and N is large?

    (c) Simulate the throughput obtained using the proportional fair scheduling algorithm for i_c values of 1, 5, and 10.

Bibliography

[1] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1992.

[2] T. S. Rappaport, *Wireless Communications–Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

[3] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel full duplex wireless communication", Proc. 16th Annual International Conference on Mobile Computer Networks, pp. 1-12, Sep. 2010.

[4] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results", Proc. Conf. Rec. 44th Asilomar Conf. Signals Syst. Comput., pp. 1558-1562, Nov. 2010.

[5] B. Smida, A. Sabharwal, G. Fodor, G. C. Alexandropoulos, H. A. Suraweera, and C. -B. Chae, "Full-duplex wireless for 6G: Progress brings new opportunities and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 9, pp. 2729-2750, Sept. 2023.

[6] G. Leus, S. Zhou, and G. B. Giannakis, "Orthogonal multiple access over time-and frequency-selective channels," *IEEE Trans. Inform. Theory*, pp. 1942–50, August 2003.

[7] S. Verdú, "Demodulation in the presence of multiuser interference: Progress and misconceptions," in D. Docampo, A. Figueiras, and F. Perez-Gonzalez, Eds., *Intelligent Methods in Signal Processing and Communications*, pp. 15–46, Birkhäuser, Boston, 1997.

[8] W. Stallings, *Wireless Communications and Networks*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 2005.

[9] B. Gundmundson, J. Sköld, and J. K. Ugland, "A comparison of CDMA and TDMA systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 732–5, May 1992.

[10] V. Khammammetti and S. K. Mohammed, "OTFS-based multiple-access in high Doppler and delay spread wireless channels," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 528-531, April 2019.

[11] Z. Ding, R. Schober, P. Fan and H. Vincent Poor, "OTFS-NOMA: An efficient approach for exploiting heterogenous user mobility profiles," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7950-7965, Nov. 2019.

[12] M. Gudmundson, "Generalized frequency hopping in mobile radio systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 788–91, May 1993.

[13] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech.*, pp. 303–12, May 1991.

[14] P. Jung, P. W. Baier, and A. Steil, "Advantages of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications," *IEEE Trans. Veh. Tech*., pp. 357–64, August 1993.

[15] J. Chuang and N. Sollenberger, "Beyond 3G: Wideband wireless data access based on OFDM and dynamic packet assignment," *IEEE Commun. Mag*., pp. 78–87, July 2000.

[16] K. R. Santhi, V. K. Srivastava, G. SenthilKumaran, and A. Butare, "Goals of true broad band's wireless next wave (4G-5G)," *Proc. IEEE Veh. Tech. Conf*., pp. 2317–21, October 2003.

[17] M. Frodigh, S. Parkvall, C. Roobol, P. Johansson, and P. Larsson, "Future-generation wireless networks," *IEEE Wireless Commun. Mag*., pp. 10–17, October 2001.

[18] E. Anderlind and J. Zander, "A traffic model for non-real-time data users in a wireless radio network," *IEEE Commun. Lett*., pp. 37–9, March 1997.

[19] K. Pahlavan and P. Krishnamurthy, *Principles of Wireless Networks*: *A Unified Approach*, Prentice-Hall, Englewood Cliffs, NJ, 2002.

[20] N. Abramson, "The ALOHA system–Another alternative for computer communications," *Proc. Amer. Federation Inform. Proc. Soc. Fall Joint Comput. Conf*., pp. 281–5, November 1970.

[21] A. Chockalingam and M. Zorzi, "Energy consumption performance of a class of access protocols for mobile data networks," *Proc. IEEE Veh. Tech. Conf*., pp. 820–4, May 1998.

[22] S. Haykin and M. Moher, *Modern Wireless Communications*, Prentice-Hall, Englewood Cliffs, NJ, 2005.

[23] P. Karn, "MACA: A new channel access method for packet radio," *Proc. Comput. Network Conf*., pp. 134–40, September 1990.

[24] V. Bharghavan, A. Demers, S. Shenkar, and L. Zhang, "MACAW: A media access protocol for wireless LANs," *Proc. ACM SIGCOMM*, vol. 1, pp. 212–25, August 1994.

[25] *IEEE Standard for Wireless LAN Medium Access Control* (*MAC*) *and Physical Layer* (*PHY*) *Specifications*, IEEE Standard 802.11, 1997.

[26] Z. J. Haas, J. Deng, and S. Tabrizi, "Collision-free medium access control scheme for ad hoc networks, *Proc. Military Commun. Conf*., pp. 276–80, 1999.

[27] S.-L. Wu, Y.-C. Tseng, and J.-P. Sheu, "Intelligent medium access for mobile ad hoc networks with busy tones and power control," *IEEE J. Sel. Areas Commun*., pp. 1647–57, September 2000.

[28] N. Abramson, "Wide-band random-access for the last mile," *IEEE Pers. Commun. Mag*., pp. 29–33, December 1996.

[29] A. Asadi and V. Mancuso. "A survey on opportunistic scheduling in wireless communications." *IEEE Communications Surveys and Tutorials*, pp. 1671-1688, January 2013.

[30] D. J. Goodman, R. A. Valenzuela, K. T. Gayliard, and B. Ramamurthi, "Packet reservation multiple access for local wireless communications," *IEEE Trans. Commun*., pp. 885–90, August 1989.

[31] N. B. Mehta and A. J. Goldsmith, "Effect of fixed and interference-induced packet error probability on PRMA," *Proc. IEEE Internat. Conf. Commun*., pp. 362–6, June 2000.

[32] P. Agrawal, "Energy efficient protocols for wireless systems," *Proc. Internat. Sympos. Pers., Indoor, Mobile Radio Commun.*, pp. 564–9, September 1998.

[33] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Tech.*, pp. 57–62, February 1992.

[34] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Tech.*, pp. 641–6, November 1993.

[35] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Trans. Commun.*, pp. 226–8, February–April 1994.

[36] E. Seneta, *Nonnegative Matrices and Markov Chains*, Springer-Verlag, New York, 1981.

[37] N. Bambos, S. C. Chen, and G. J. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," $IEEE/ACM$ *Trans. Network.*, pp. 583–97, October 2000.

[38] S. Kandukuri and N. Bambos, "Power controlled multiple access (PCMA) in wireless communication networks," *Proc. IEEE Infocom Conf.*, pp. 386–95, March 2000.

[39] P. P. Bergmans and T. M. Cover, "Cooperative broadcasting," *IEEE Trans. Inform. Theory*, pp. 317–24, May 1974.

[40] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, pp. 1083–1102, March 2001.

[41] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[42] L.-F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, pp. 1439–49, October 1993.

[43] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, pp. 279–80, March 1974.

[44] R. Pickholtz, L. Milstein, and D. Schilling, "Spread spectrum for mobile communications," *IEEE Trans. Veh. Tech.*, pp. 313–22, May 1991.

[45] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[46] D. Tse and S. Hanly, "Multiaccess fading channels – Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, pp. 2796–2815, November 1998.

[47] S. Hanly and D. Tse, "Multiaccess fading channels – Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, pp. 2816–31, November 1998.

[48] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels – Part II: Outage capacity," *IEEE Trans. Inform. Theory*, pp. 1103–27, March 2001.

[49] L. Li, N. Jindal, and A. J. Goldsmith, "Outage capacities and optimal power allocation for fading multiple access channels," *IEEE Trans. Inform. Theory*, pp. 1326–47, Apri1 2005.

[50] N. Jindal and A. J. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," *IEEE Trans. Inform. Theory*, pp. 2895–2909, November 2003.

[51] E. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.

[52] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, pp. 1691–1706, July 2003.

[53] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," *Proc. IEEE Globecom Conf.*, pp. 1344–8, November 2001.

[54] H. Weingarten, Y. Steinberg, and S. Shamai, " The capacity region of the Gaussian MIMO broadcast channel," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 174, June 2004.

[55] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, pp. 439–41, May 1983.

[56] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *Proc. Internat. Sympos. Inform. Theory Appl.*, pp. 681–4, November 2000.

[57] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, pp. 364–75, March 1996.

[58] D. Tse and P. Viswanath, *Foundations of Wireless Communications*, Cambridge University Press, 2005.

[59] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," *Proc. IEEE Internat. Conf. Commun.*, pp. 331–5, June 1995.

[60] N. Jindal, S. Vishwanath, and A. J. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inform. Theory*, pp. 768–83, May 2004.

[61] S. Verdú, "Multiple-access channels with memory with and without frame synchronism," *IEEE Trans. Inform. Theory*, pp. 605–19, May 1989.

[62] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecommun.*, pp. 585–96, November 1999.

[63] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for vector multiple access channels," *IEEE Trans. Inform. Theory*, pp. 145–52, January 2004.

[64] R. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, pp. 773–85, May 1993.

[65] A. J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inform. Theory*, pp. 219–40, January 2001.

[66] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, pp. 2658–68, October 2003.

[67] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, pp. 1912–21, August 2003.

[68] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. J. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, pp. 1570–9, Apri1 2005.

[69] P. Vishwanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, pp. 1277–94, June 2002.

[70] TIA/EIA IS-856, "CDMA 2000: High rate packet data air interface specification," November 2000.

[71] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency–high data rate personal communication wireless system," *Proc. IEEE Veh. Tech. Conf.*, pp. 1854–8, May 2000.

[72] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, and P. Whiting, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, pp. 150–4, February 2001.

[73] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE J. Sel. Areas Commun.*, pp. 2053–64, October 2001.

[74] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing trade-off in multiple-access channels," *IEEE Trans. Inform. Theory*, pp. 1859–74, September 2004.

[75] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, pp. 684–701, June 2003.

[76] M. Sharif and B. Hassibi, "Scaling laws of sum rate using time-sharing, DPC, and beamforming for MIMO broadcast channels," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 175, June 2004.

[77] T. Yoo and A. J. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," *Proc. IEEE Internat. Conf. Commun.*, May 2005.

[78] R. W. Heath, Jr., M. Airy, and A. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," *Proc. Asilomar Conf. Signals, Syst., Comput.*, pp. 1194–9, November 2001.

[79] N. Jindal and A. Goldsmith, "DPC vs. TDMA for MIMO broadcast channels," *IEEE Trans. Inform. Theory*, pp. 1783–94, May 2005.

[80] N. Al-Dhahir, C. Fragouli, A. Stamoulis, W. Younis, and R. Calderbank, "Space-time processing for broadband wireless access," *IEEE Commun. Mag.*, pp. 136–42, September 2002.

[81] M. Brehler and M. K. Varanasi, "Optimum receivers and low-dimensional spreaded modulation for multiuser space-time communications," *IEEE Trans. Inform. Theory*, pp. 901–18, April 2003.

[82] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, "Multiuser joint equalization and decoding of space-time codes," *Proc. IEEE Internat. Conf. Commun.*, pp. 2643–7, May 2003.

[83] H. Dai and H. V. Poor, "Iterative space-time processing for multiuser detection in multipath CDMA channels," *IEEE Trans. Signal Proc.*, pp. 2116–27, September 2002.

[84] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, "On interference cancellation and high-rate space-time codes," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 238, June 2003.

[85] S. J. Grant and J. K. Cavers, "System-wide capacity increase for narrowband cellular systems through multiuser detection and base station diversity arrays," *IEEE Trans. Wireless Commun.*, pp. 2072–82, November 2004.

[86] Z. Liu and G. B. Giannakis, "Space-time block-coded multiple access through frequency-selective fading channels," *IEEE Trans. Commun.*, pp. 1033–44, June 2001.

[87] K.-K. Wong, R. D. Murch, and K. B. Letaief, “Performance enhancement of multiuser MIMO wireless communication systems,” *IEEE Trans. Commun.*, pp. 1960–70, December 2002.

Chapter 15

Cellular Networks

Cellular networks are infrastructure-based wireless networks that divide the geographic region of coverage into cells, as shown in Figure 15.1. Channels, or set of channels, C_i are reused at spatially-separated cells to make efficient use of spectrum. A base station is located within each cell. Under ideal propagation conditions, users within a given cell communicate with the base station in that cell, although in practice the choice is based on the SINR between the user and the base station. The channels from a base station to all users in its cell define the downlink of the cell, and the channels from all users in a cell to the cell base station define the uplink of the cell.

As described in Section 1.2.2, all base stations in a given region, such as those shown in Figure 15.1, are connected to a mobile telephone switching office (MTSO). The MTSO is the central controller across all its connected base stations; it coordinates user authentication, access and handoff, transmission scheduling, dynamic channel allocation, as well as routing to the Public Switched Telephone Network (PSTN) and the Internet. A handoff between base stations is initiated by the MTSO when the signal quality of a user to its base station decreases below a given threshold. This occurs when a user moves between cells and can also be caused by fading or shadowing within a cell. If no neighboring base station has available channels of acceptable quality then the handoff attempt fails and the call will be dropped. The infrastructure associated with a cellular network consisting of base stations, MTSOs, and connectivity to the PSTN and Internet is in contrast to ad hoc wireless networks, described in Chapter 16, which have no such infrastructure.

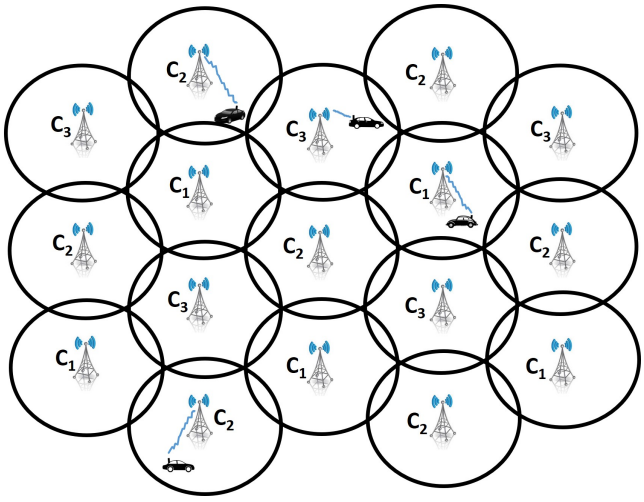


Figure 15.1: Cellular Network with Channel Reuse

Most cellular networks are designed so that users transmit directly to a base station, with no multihop routing through intermediate wireless nodes. In general these single-hop routes have lower delay and loss, higher data

rates, and more flexibility than multihop routes. For these reasons, the performance of cellular networks tends to be much better than in networks without infrastructure. However, it is sometimes more expensive or simply not feasible or practical to deploy the infrastructure associated with a cellular network in a given environment, in which case ad hoc wireless networks are the best option despite their typically inferior performance.

The focus of this chapter is on cellular network design, analysis, and fundamental performance limits. We will first describe the basic principles in cellular network design along with interference mitigation techniques to improve their performance. The capacity of a cellular network is then derived for both orthogonal and non-orthogonal channelization methods. We next discuss dynamic allocation of channels and power to optimize system performance as user locations and propagation conditions change. The chapter closes with an analysis of the fundamental rate limits of cellular networks in terms of their Shannon capacity and area spectral efficiency.

15.1 Design Principles

15.1.1 Channel Reuse and SINR

The underlying design principles of cellular networks are based on exploiting the power falloff with distance of signal propagation in order to reuse the same channel at spatially separated locations. The signaling dimensions associated with the spectrum of the system is channelized using one of the orthogonal or nonorthogonal techniques discussed in Section 14.2. We will focus on time-division via TDMA, frequency-division via FDMA, and time-frequency block division via OFDMA for orthogonal channelization. Non-orthogonal CDMA will be used to obtain nonorthogonal channelization; orthogonal CDMA will not be considered in this chapter. These techniques are sometimes combined with space-division channelization as well. Each channel obtained by channelizing the signaling dimensions is assigned to a unique *channel set* C_n. Every cell is then assigned a channel set, and these sets can be reused at spatially separated locations, as Figure 15.1 illustrates. This reuse of channels is called *frequency reuse* or *channel reuse*. Cells that are assigned the same channel set, called *co-channel cells*, must be spaced far enough apart that interference between users in the cells does not degrade signal quality below the level needed for minimum acceptable performance. The required spacing of co-channel cells to achieve acceptable signal quality depends on the channelization technique, the signal propagation characteristics, and the desired performance for each user.

Uplink and downlink design for a single cell was described in Section 14.2, and many of the same design principles and analyses apply to cellular networks once appropriately modified to include the impact of channel reuse. The trade-offs associated with different channelization techniques are different in cellular networks than in a single cell, since each technique must cope with interference from outside its cell, known as *intercell* or *co-channel* interference. In addition, systems with nonorthogonal channelization must also deal with interference from within a cell, called *intracell* interference. This intracell interference also arises in systems with orthogonal channelization when multipath, synchronization errors, and other practical impairments compromise the orthogonality.

Although non-orthogonal CDMA has both intracell and intercell interference inherent to its design, all interference is attenuated by the code cross-correlation. In contrast, orthogonal multiple access techniques have no intracell interference under ideal operating conditions. However, their intercell interference has no reduction from processing gain as in spread spectrum systems. The amount of both intercell and intracell interference experienced by a given user is captured by its signal-to-interference-plus-noise power ratio, defined as

$$\text{SINR} = \frac{P_r}{N_0 B + P_I}, \tag{15.1}$$

where P_r is the received signal power and P_I is the received power associated with both intracell and intercell interference. In CDMA systems, P_I is the interference power after de-spreading. We typically compute the bit

error rate of a mobile based on SINR rather than SNR, although this approximation is not precisely accurate if the random interference cannot be characterized by Gaussian statistics.

A larger intercell interference reduces SINR and hence increases user BER. Intercell interference can be kept small by separating cells operating on the same channel by a large distance. However, the number of users that a system can accommodate is maximized by reusing frequencies as often as possible. Thus, the best cellular system design places users that share the same channel at a separation distance where the intercell interference is just below the maximum tolerable level for the required data rate and BER. Good cellular system designs are *interference limited*, meaning that the interference power is much larger than the noise power. Therefore, noise is generally neglected in the study of these systems. In this case SINR reduces to the signal-to-interference power ratio $\mathrm{SIR} = P_r/P_I$. In interference-limited systems, since the BER of users is determined by SIR, the number of users that can be accommodated is limited by the interference they cause to other users.

15.1.2 Interference Reduction Techniques

Techniques to reduce interference increase the SIR and therefore increase the number of users the system can accommodate for a given BER constraint. The main techniques to reduce interference include signal processing techniques such as multiuser detection (MUD), interference precancellation, and frequency-hopping overlays, as well as base station antenna design techniques, which are covered in Section 15.2. Multiuser detectors jointly detect the desired signal and some or all interference, so that the detected interference can be mitigated or cancelled out. There are trade-offs between performance and complexity with different multiuser detection methods, especially when there are a large number of interferers. Multiuser detection methods were described in Section 13.4 for the uplink of a single-cell CDMA system. However, these detection methods can also be applied to intercell interference signals – both at the base station [39], where processing complexity is less of a constraint, and in the mobile device to cancel a few dominant interferers [40].

Interference precancellation takes advantage of the fact that, in the downlink, the base station has knowledge of (i) interference between users within its cell and (ii) the interference its transmission causes to mobiles in other cells. This knowledge can be utilized in the base station transmission to presubtract interference between users [41, 42, 43]. Interference presubtraction by the base station has its roots in the capacity-achieving strategy for downlinks, which uses a "dirty paper coding" transmission technique (see Section 14.6.5) for interference presubtraction developed by Costa [44]. Numerical results in [41, 42, 43] indicate that interference presubtraction can lead to an order-of-magnitude capacity increase in cellular systems with a large number of base station antennas. However, the presubtraction requires channel CSI at the transmitter–in contrast to multiuser detection, which does not require transmitter CSI. Interference cancellation is a special case of Base Station Cooperation (BSC), whereby multiple base stations within a network share information and coordinate their transmitter and receiver protocols to improve the performance for all their users. Cooperation techniques can include MIMO antenna parameter optimization, power control, channel assignment, as well as general real-time coordination of downlink transmission and uplink reception across the antennas associated with the cooperating base stations. BSC generally leads to reduced interference between co-channel cells as well as the distributed MIMO benefits of increased capacity and diversity. While BSC was incorporated into the 4G cellular standard, the relatively modest performance gains that were obtained did not justify the significantly increased complexity associated with base station cooperation. For that reason BSC is not part of current cellular or WiFi standards.

Intercell interference in the uplink is often dominated by one or two mobile users located near the cell boundaries closest to the base station serving the desired user. In TDMA or FDMA systems, the impact of these worst-case interferers can be mitigated by superimposing frequency hopping (FH) on top of the TDMA or FDMA channelization. With this FH overlay, all mobiles change their carrier frequency according to a unique hopping pattern. Since the hopping pattern of the worst-case interferers differs from that of the desired user, these interferers would cause intercell interference to the desired signal only when the two hop patterns overlapped in both time and fre-

quency, which seldom occurs. Thus, the FH overlay has the effect of interference averaging, causing intercell interference from any given cell to be averaged relative to all interferer locations. This greatly reduces the effect of interference associated with mobiles on cell boundaries. For this reason, an FH overlay is used in the GSM cellular system.

15.1.3 Cell Shape and Size

The cell shape for a given base station in a cellular network is determined by the contours of constant received power associated with its transmission at the cell boundary, as was discussed in Section 2.7.3. The cell shapes depicted in Figure 15.1 are circles, which assumes the simplified (or free-space) path loss model. Under this propagation model, a receiver at any fixed distance from the transmitting base station will have the same received power, leading to a circular contour of constant received power. In general, given the large variation in path loss, shadowing, and multipath fading for different environments, the contours of constant received power associated with any given base station is a highly irregular shape. To simplify cellular network design and analysis, models of these systems typically assume regular cell shapes that are also *tessellating*. Tessellating cell shapes can be laid next to each other with no overlap and so cover the entire geographical region of a cellular network without any gaps. Tessellating cell shapes include hexagons, rectangles, squares, diamonds, and triangles. When propagation in a cellular network follows the simplified path-loss model leading to circular contours of constant power, a hexagon provides a reasonable approximation for cell shape that is also tessallating. Hexagons are commonly used to approximate cell shapes for cellular systems where base stations are placed on towers or at the top of a building with coverage areas on the order of a few square kilometers. When base stations are placed closer to the ground and transmit at lower power for coverage over tens to hundreds of meters, diamonds as tessellating cell shapes tend to better approximate the contours of constant power, especially in typical urban street grids [2, 3]. Very small cells and indoor cells are heavily dependent on the propagation environment, making it difficult to accurately approximate contours of constant power using a tessellating shape [4].

While cell shape is dictated by the propagation environment of a cellular network's base stations, cell size is a design choice determined by the base stations' transmit power and the desired received power at the cell boundary. Since signal power falls off relative to distance, reducing the transmit power will reduce the distance at which a base station's signal is received at the desired power level. The number of users that can be accommodated within a given system can be increased by shrinking the size of a cell, as long as all aspects of the system scale so that the SINR of each user remains the same. Specifically, consider the large and small cells shown in Figure 15.2. Suppose the large cell in this figure represents one cell in a cellular system, where each cell accommodates K users. If the cell size is shrunk to the smaller cell size shown in Figure 15.2 (by reducing transmit power) and if everything in the system (including propagation loss) scales so that the SINR in the small cells is the same as in the original large cell, then this smaller cell can also accommodate K users. Since there are 19 small cells within the large cell, the new system with smaller cells can accommodate $19 \times K$ users within the area of one large cell – a 19-fold increase in user capacity. However, propagation characteristics typically change as cell size shrinks, so the system does not scale perfectly. Moreover, a smaller cell size increases the rate at which handoffs occur, which increases the dropping probability if the percentage of failed hand-offs stays the same. Infrastructure and deployment costs also increase for systems with smaller cells, since more base stations must be deployed per unit area and each base station site requires site rental or purchase, power, and backbone network connectivity. Thus, although smaller cells generally increase capacity, they also have their disadvantages. To exploit the benefits of both large and small cells, most cellular networks embed small cells within the geographical area of a large cell that experiences high traffic, leading to a heterogeneous network (Hetnet) architecture. This allows the high capacity of a small cell to serve high-traffic areas, while large cells provide uniform coverage throughout the region. Cellular hetnet architectures will be discussed in Section 15.5.

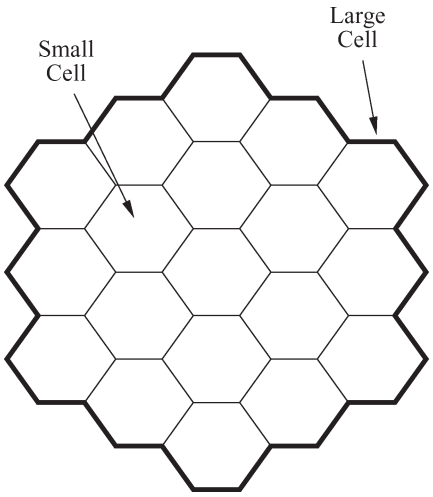


Figure 15.2: Capacity increase by shrinking cell size.

15.1.4 Channel Allocation and Reuse Distance

Channel allocation for each cell is a key element of cellular system design; it determines how much interference is experienced by users occupying the same channel as well as the spectral efficiency of the system overall. Channel allocation can be fixed based on average propagation conditions for a given system layout, or it can be dynamic based on real-time measurements of signal and interference power for each user within a cell. This section focuses on fixed channel allocation for each cell to illustrate the concept of reuse distance and how user SINR depends on this distance. Dynamic channel allocation, described in Section 15.4.2, uses real-time optimization based on varying channel conditions and user locations to determine which channel to assign to each user. Ideally the optimal dynamic channel allocation across all users minimizes co-channel interference such that the same channel can be used in every cell, thereby maximizing spectral efficiency.

The design of a channel allocation scheme and the associated channel reuse is generally different for systems with orthogonal channels such as TDMA, FDMA, and OFDMA versus those with nonorthogonal CDMA channels. In particular, cellular systems with orthogonal channels have no intracell interference under ideal conditions. However, in systems with fixed channel reuse, cells using the same channel set are typically spaced several cells away, since co-channel interference from adjacent cells can be very large. In contrast, non-orthogonal CDMA systems exhibit both intercell and intracell interference, but all interference is attenuated by the cross-correlation of the spreading codes. Due to this interference reduction, non-orthogonal CDMA systems typically reuse channel sets in every cell, which will be our assumption throughout this chapter. The simplicity of reusing channels in every cell greatly reduces the complexity of CDMA cellular system design and operation. This was one of the major reasons for its adoption in the 3G cellular standard.

The basic premise of fixed channel allocation for cellular systems with orthogonal channels consists of cell clustering and channel assignment. In interference-limited systems, each user's BER is based on his received SIR: the ratio of his received signal power to his intracell and intercell interference power. The received signal powers associated with the desired signal, the intercell interference, and the intracell interference are determined by the characteristics of the channel between the desired or interfering transmitters and the desired receiver. The average SIR is normally computed based on path loss alone, with median shadowing attenuation incorporated into the path-loss models for the signal and interference. Random variations due to shadowing and flat fading are then treated as statistical variations about the path loss.

Since path loss is a function of propagation distance, the reuse distance D between cells using the same

channel is an important parameter in determining average intercell interference power. Reuse distance is defined as the distance between the centers of cells that use the same channel set. It is a function of cell shape, cell size, and the number of intermediate cells between the two cells sharing the same channel set. Given a required average SINR for a particular performance level, we can find the corresponding minimum reuse distance that meets this performance target. The focus of this section is planning the cellular system layout based on a minimum reuse distance requirement.

Figure 15.3 illustrates the reuse distance associated with a given channel reuse pattern for hexagonal and diamond-shaped cells. Cells that are assigned the same channel set C_n are indicated in the figure. This pattern of channel reuse for both cell shapes is based on the notion of cell clustering, discussed in more detail below. The reuse distance D between these cells is the minimum distance between the dots at the center of cells using channel set C_n. The radius R of a cell is also shown in the figure. For hexagonal cells, R is defined as the distance from the center of a cell to a vertex of the hexagon. For diamond-shaped cells, R is the distance from the cell center to the middle of a side.

For diamond-shaped cells it is straightforward to compute the reuse distance D based on the number of intermediate cells N_I between co-channel cells and the cell radius R. Specifically, the distance across a cell is $2R$. The distance from a cell center to its boundary is R and the distance across the N_I intermediate cells between the co-channel cells is $2RN_I$. Thus, $D = R + 2RN_I + R = 2R(N_I + 1)$.

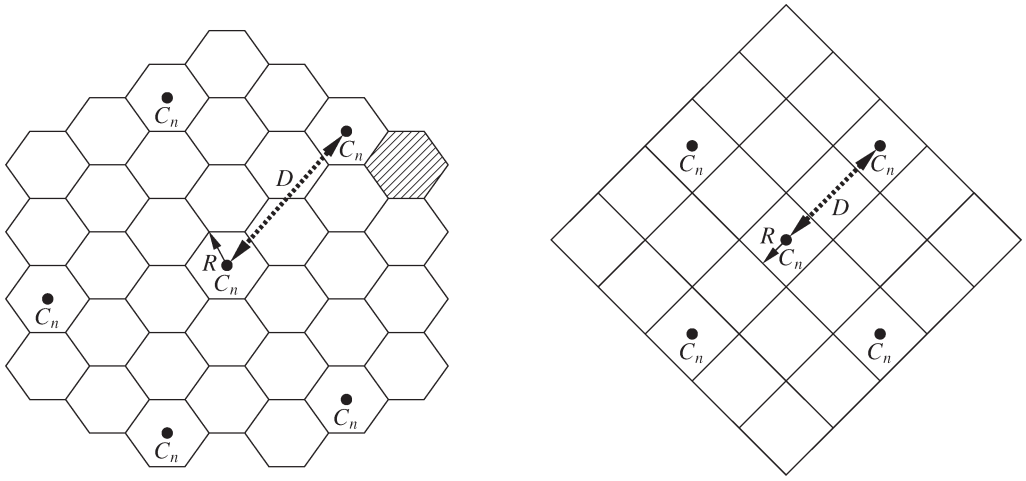


Figure 15.3: Reuse distance D with a given reuse pattern for hexagonal and diamond-shaped cells.

Reuse distance for hexagonally shaped cells is more complicated to determine, since there might not be an integer number of cells between two co-channel cells. In the left side of Figure 15.3, if channel C_n is used in the center cell and again in the shaded cell, then there are two cells between co-channel cells and the reuse distance is easy to find. However, if C_n is reused in the cell adjacent to the shaded cell, as is shown in the figure, then there is not an integer number of cells separating the co-channel cells. The assignment shown in Figure 15.3 is needed to create cell clusters, as discussed in more detail below.

The procedure for channel assignment in hexagonal cells is as follows. Consider the cell diagram in Figure 15.4, where again R is the hexagonal cell radius. Denote the location of each cell by the pair (i, j). Then, assuming cell A to be centered at the origin $(0, 0)$, the location of the cell at (i, j) relative to cell A is obtained by moving i cells along the u-axis, turning 60° counterclockwise, and moving j cells along the v-axis. For example, cell G is located at $(0, 1)$, cell S at (1,1), cell P at $(-2, 2)$, and cell M at $(-1, -1)$. It is straightforward to show that the distance between cell centers of adjacent cells is $\sqrt{3}R$, and the distance between the centers of the cell located at the point (i, j) and cell A located at $(0, 0)$) is given by

$$D = \sqrt{3}R\sqrt{i^2 + j^2 + ij}. \tag{15.2}$$

Example 15.1: Find the reuse distance D for the channel reuse pattern illustrated in Figure 15.3 for both the diamond- and hexagonally shaped cells as a function of cell radius R.
Solution: For the diamond-shaped cells, there is $N_I = 1$ cell between co-channel cells. Thus, $D = 2R(N_I + 1) = 4R$. For the hexagonal cells shown in Figure 15.3, the reuse pattern moves two cells along the u-axis and then one cell along the v-axis. Thus, $D = \sqrt{3}R\sqrt{2^2 + 1^2 + 2} = 4.58R$.

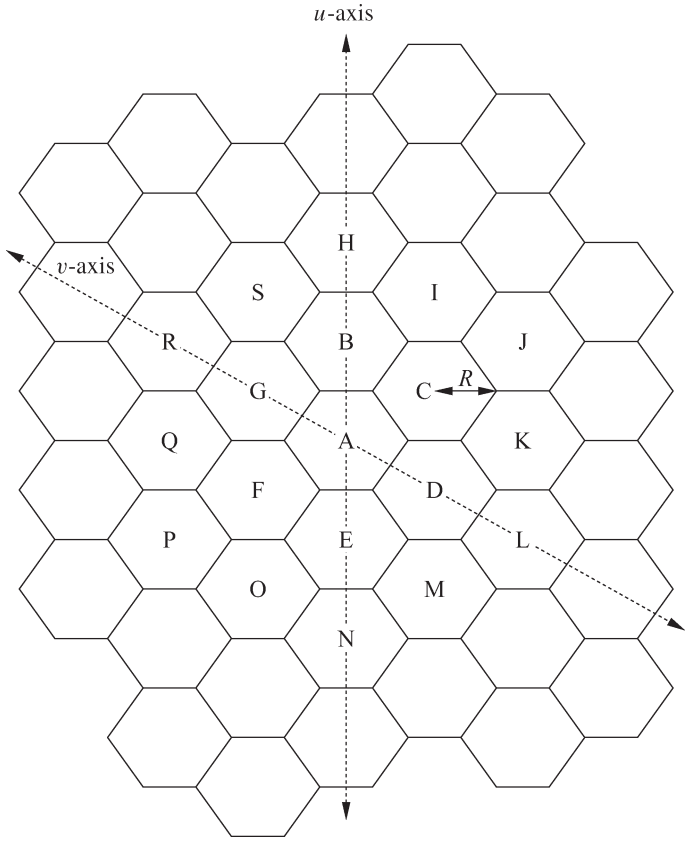


Figure 15.4: Axes for reuse distance in hexagonal cells.

Given a minimum acceptable reuse distance $D_{\min}$, we would like to maintain this distance throughout the cell grid while reusing channels as often as possible. This requires spatially repeating the *cell clusters* for channel assignment, where each cell in the cluster is assigned a unique channel set that is not assigned to any other cell in the cluster. In order to spatially repeat, cell clusters must tessellate. For diamond-shaped cells, a tessellating cell cluster forms another diamond, with K cells on each side. The channel set assigned to the nth cell in the cluster is denoted by C_n, $n = 1, \ldots, N$, where N is the number of unique channel sets, and the pattern of channel assignment is repeated in each cluster. This ensures that cells using the same channel are separated by a reuse distance of at least $D = 2KR$. The number of cells per cluster is $N = K^2$, which is also called the *reuse factor*: since $D = 2KR$, we have $N = .25(D/R)^2$. If we let N_c denote the number of channels per cell (i.e., the number of channels in C_n) and N_T the total number of channels, then $N = N_T/N_c$. A small value of N indicates efficient

channel reuse (channels reused more often within a given area for a fixed cell size and shape). However, a small N also implies a small reuse distance, since $D = 2KR = 2\sqrt{N}R$, which can lead to substantial intercell interference.

For hexagonal cells, we form cell clusters through the following iterative process. The total bandwidth is first broken into N channel sets $C_1, \ldots, C_N$, where N is the cluster size. We assign the first channel set C_1 to any arbitrary cell. From this cell, we move i cells along a chain of hexagons in any direction, turn counterclockwise by 60°, and move j cells along the hexagon chain in this new direction; channel set C_1 is then assigned to this jth cell. Going back to the original cell, we repeat the process in a different direction until we have covered all directions starting from the initial cell. This process is shown in Figure 15.5 for $i = 3$ and $j = 2$. To assign channel set C_1 throughout the region, we repeat the iterative process starting from one of the cells assigned channel set C_1 in a prior iteration until no new assignments can be made starting from any location that has been assigned channel set C_1. Then a new cell that has not been assigned any channel set is selected and assigned channel set C_2, and the iterative channel assignment process for this new channel set is performed. The process is repeated until all cells have been assigned a unique channel set, which results in cell clusters as illustrated in Figure 15.6. The reuse distance between all cells using the same channel set is then $D = \sqrt{3}R\sqrt{i^2 + j^2 + ij}$. We can obtain the approximate cluster size associated with this process by finding the ratio of cluster area to cell area. Specifically, the area of a hexagonal cell is $A_{\text{cell}} = 3\sqrt{3}R^2/2$, and the area of a cluster of hexagonal cells is $A_{\text{cluster}} = \sqrt{3}D^2/2$. Hence, the number of cells per cluster is

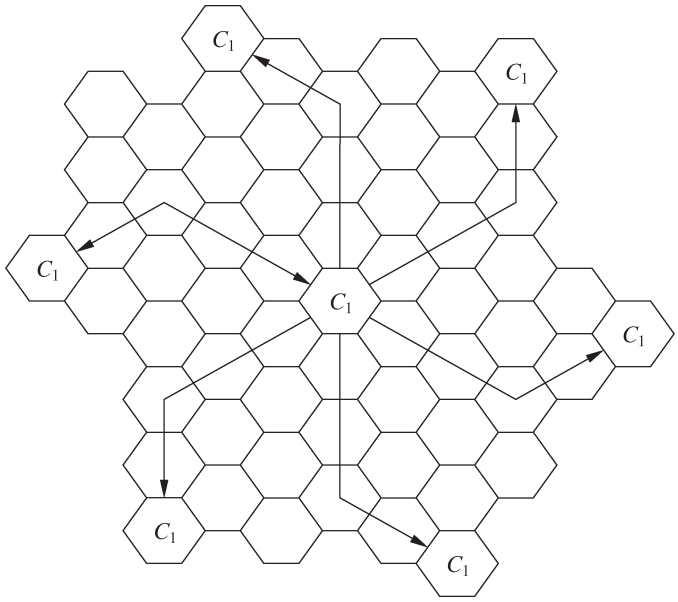


Figure 15.5: Channel assignment in hexagonal cells.

$$N = \frac{A_{\text{cluster}}}{A_{\text{cell}}} = \frac{\sqrt{3}D^2/2}{3\sqrt{3}R^2/2} = \frac{1}{3}\left(\frac{D}{R}\right)^2 = \frac{1}{3}\left(\frac{3R^2(i^2 + j^2 + ij)}{R^2}\right) = i^2 + j^2 + ij. \tag{15.3}$$

As with diamond cells, a small N indicates more efficient channel reuse but also a smaller reuse distance $D = R\sqrt{3N}$, leading to more intercell interference.

15.1.5 Base Station Antenna Design

In this section we describe base station antenna design techniques and their impact on interference reduction and capacity increase. A common technique to reduce interference via base station antenna design is sectorization.

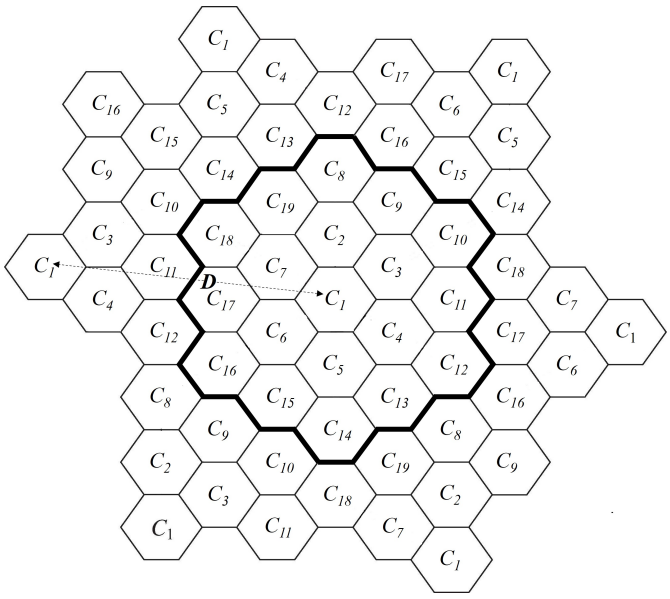


Figure 15.6: Cell clusters for hexagonal cells.

As described in Section 2.9, antenna sectorization uses directional antennas to divide up a base station's 360° omnidirectional antenna transmission into N sectors, as shown in Figure 15.7 for $N = 8$. As the figure indicates, intracell and intercell interference to a given mobile comes primarily from within its sector. Thus, sectorization reduces interference power by roughly a factor of N under heavy loading (interferers in every sector). Handoff is required as mobiles move between sectors, which increases overhead and the probability of call dropping.

Most cellular networks employ multiple-antenna arrays at the base stations in order to exploit the mulituser MIMO methods described in Section 14.10. These base station antenna arrays form narrow beams that provide high gain to the desired user's signal, and/or can create spatial nulls in the direction of interference signals [17]. However, antenna arrays can also be used for multiplexing gain, which leads to higher data rates, or diversity gain, which improves reliability. These fundamental trade-offs are described in Section 14.10. The use of multiple antennas at a given cellular base station can be dynamically adjusted to provide interference reduction, higher data rates, or diversity gain to optimize a given performance metric across all users within the cell.

Distributed antenna systems (DAS) are another technique that can be deployed within a cell to improve performance [18]. In this setting, rather than have all the base station antennas co-located with the base station, some or all of the base station antennas can be distributed throughout the cell to move them closer to the users. The distributed antennas are typically connected to their base station by high-speed reliable communication links. By having antennas closer to the users, the required transmit power between a user and its closest base station antenna is reduced, thereby reducing power consumption and interference to other users. Coverage holes due to shadowing can also be removed by placing a distributed antenna in or near the area of poor coverage. For these reasons, DAS is often utilized inside large buildings such as offices, stadiums, airports, and shopping malls to improve cellular coverage and performance.

The benefits of DAS to reduce energy consumption, reduce interference, and fill coverage holes can also be obtained through multihop routing. Multihop routing in a cellular system entails that the transmission from a base station to a user, or from a user to a base station, is routed through an intermediate node. That node may be another user, or a fixed relay node that is deployed within the cell for that purpose. While multihop routing has similar performance advantages as DAS in principle, the gains obtained in practice have not been sufficient to warrant its

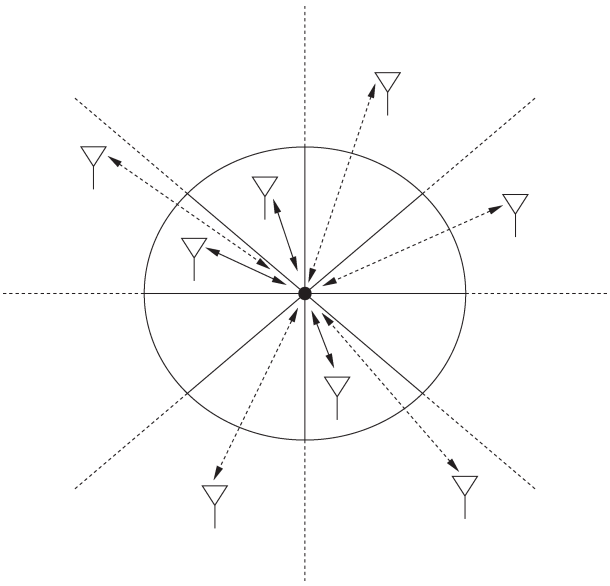


Figure 15.7: Circular cell sectorization for $N = 8$.

widespread use.

15.2 SIR and User Capacity

In this section we analyze the SIR of users in a cellular system and the number of users per cell that can be supported for a given SIR target. We neglect the impact of noise on performance under the assumption that the system is interference-limited, although the calculation can be easily extended to include noise. The SIR for a given user in a cellular system is a function of the transmit power associated with the user and its interferers, the random channel between the user and its base station, and the random channel between each interfering transmitter and the user's receiver. Thus, the SIR is the ratio of two random numbers whose distributions depend on user and base station locations and antenna characteristics as well as propagation characteristics. There are generally three approaches to obtain the SIR distribution and associated performance analysis in a cellular system: extensive field measurements [19], simulations that average over variations in the base station and user locations as well as the user and interferer channels [20], or a simplified mathematical model that trades accuracy in exchange for analytical tractability. In this section we develop a tractable mathematical model for SIR and user capacity in both orthogonal and non-orthogonal systems based on the simplified path loss model for propagation. The model can be extended to more complex propagation models for path loss, shadowing, and fading; for these models numerical solutions or approximations are needed to obtain the SIR distribution (see e.g. [21, 22]). An alternative analytical approach using stochastic geometry that models base station locations as Poisson point processes coupled with fading and shadowing on each link is described in [25, 26].

Our SIR model assumes the simplified path-loss model (2.24) with reference distance $d_0 = 1$ meter for our path-loss calculations, so $P_r = P_t L d^{-\gamma}$, where L is a constant equal to the average path loss at $d = d_0$ and γ is the path-loss exponent. Since L is typically the same for both the desired signal and interference, its value cancels out in the SIR power ratio; hence we set $L = 1$ in the analysis that follows. The path-loss exponent associated with in-cell signal propagation is denoted by γ_I, while γ_0 denotes the path-loss exponent associated with out-of-cell (intercell) signal propagation. These path-loss exponents may be different, depending on the propagation environment and cell size [6]. Using the simplified path-loss model, we will derive expressions for SIR under

both orthogonal and nonorthogonal access techniques. We then find user capacity for each model, defined as the maximum number of users per cell that the system can support without violating an SIR target value.

The SIR of a signal is typically used to compute the BER performance associated with that signal. Specifically, the interference is approximated as additive white Gaussian noise (AWGN) and then formulas for the BER versus SNR are applied. For example, from (6.8), performance of uncoded BPSK without fading yields $P_b = Q(\sqrt{2 \cdot \text{SIR}})$, and from (6.69), performance when the desired signal exhibits Rayleigh fading yields $\bar{P}_b \approx .25/\text{SIR}$ for high SIRs. Although we have assumed the simplified path-loss model in the SIR formulas that follow, path-loss models of greater complexity can also be incorporated for a more accurate SIR approximation. However, there are a number of inaccuracies in the model that are not easy to fix. In particular, approximating the interference as Gaussian noise is accurate for a large number of interferers, as is the case for CDMA systems, but not accurate for a small number of interferers, as in TDMA and FDMA systems. Moreover, the performance computation in fading neglects the fact that the interferers also exhibit fading, which results in a received SIR that is the ratio of two random variables. This ratio has a complex distribution that is not well approximated by a Rayleigh distribution or by any other common fading distribution [7]. The complexity of modeling average SIR as well as its distribution under accurate path loss, shadowing, and multipath fading models can be prohibitive. Thus, the SIR distribution is often obtained via simulations [6].

15.2.1 Orthogonal Systems (TDMA/FDMA)

In this section we compute the SIR and user capacity for cellular systems using orthogonal multiple access techniques. In these systems there is no intracell interference, so the SIR is determined from the received signal power and the interference resulting from co-channel cells. Under the simplified path-loss model, the received signal power for a mobile located at distance d from its base station on both the uplink and the downlink is $P_r = P_t d^{-\gamma_I}$. The average intercell interference power is a function of the number of out-of-cell interferers. For simplicity we will neglect interference from outside the first ring of M interfering cells. This approximation is accurate when the path-loss exponent γ_0 is relatively large, since then subsequent interfering rings have a much larger path loss than the first ring. We assume that all transmitters send at the same power P_t: the impact of power control will be discussed in Section 15.3.2. Let us assume the user is at distance $d < R$ and there are M interferers at distance d_i, $i = 1, \ldots, M$, from the intended receiver (located at the mobile in the downlink and the base station in the uplink). The resulting SIR is then

$$\text{SIR} = \frac{d^{-\gamma_I}}{\sum_{i=1}^{M} d_i^{-\gamma_O}}. \tag{15.4}$$

This SIR is the ratio of two random variables, whose distribution can be quite complex. However, the statistics of the interference in the denominator has been characterized for several propagation models, and when these interferers are log-normal, their sum is also log-normal [8, Chap. 3]. Note that in general the average SIR for the uplink and downlink may be roughly the same; but the SIR for the uplink, where interferers can all be on the cell boundary closest to the base station they interfere with, generally has a smaller worst-case value than for the downlink, where interference comes from base stations at the cell centers.

The SIR expression can be simplified if we assume that the mobile is on its cell boundary, $d = R$, and all interferers are at the reuse distance D from the intended receiver. Under these assumptions the SIR reduces to

$$\text{SIR} = \frac{R^{-\gamma_I}}{MD^{-\gamma_O}}, \tag{15.5}$$

and if $\gamma_I = \gamma_O = \gamma$ then this simplifies further to

$$\text{SIR} = \frac{1}{M}\left(\frac{D}{R}\right)^{\gamma}. \tag{15.6}$$

Since D/R is a function of the reuse factor N for most cell shapes, this allows us to express the SIR in terms of N. In particular, for diamond-shaped cells, a cell cluster will be surrounded by eight adjacent clusters. Thus we have $M = 8$ and $D/R = \sqrt{4N}$. Substituting these into (15.6) yields SIR $= .125(4N)^{\gamma/2}$. From Figure 15.6, for hexagonally shaped cells we have $M = 6$ and $D/R = \sqrt{3N}$, which yields SIR $= .167(3N)^{\gamma/2}$. Both of these SIR values can be expressed as

$$\text{SIR} = a_1(a_2 N)^{\gamma/2}, \tag{15.7}$$

where $a_1 = .125$ and $a_2 = 4$ for diamond-shaped cells and where $a_1 = .167$ and $a_2 = 3$ for hexagonally shaped cells. This formula provides a simple approximation for the reuse factor required to achieve a given performance. Specifically, given a target SIR value SIR_0 required for a target BER, we can invert (15.7) to obtain the minimum reuse factor that achieves this SIR target as

$$N \geq \frac{1}{a_2}\left(\frac{\text{SIR}_0}{a_1}\right)^{2/\gamma}. \tag{15.8}$$

The corresponding minimum reuse distance D (for a given cell shape) is then calculated based on N and the cell radius R. For path-loss exponent $\gamma = 2$, (15.8) simplifies to $N \geq \text{SIR}_0/(a_1 a_2)$. If the signal has shadow fading then the analysis is more complex, but we can still generally obtain the minimum reuse factor in terms of the SIR requirement subject to some outage probability [9].

The user capacity C_u is defined as the total number of active users per cell that the system can support while meeting a common SIR target for all users. To meet this target in an orthogonal system, the reuse factor N must satisfy (15.8). Then $C_u = N_c$, where N_c is the number of channels assigned to any given cell assuming reuse factor N. The total number of orthogonal channels of bandwidth B_s that can be created from a total system bandwidth of B is $N_T = B/B_s$. Since in orthogonal systems the reuse factor N satisfies $N = N_T/N_c$, this implies that

$$C_u = \frac{N_T}{N} = \frac{B}{NB_s} = \frac{G}{N}, \tag{15.9}$$

where $G = B/B_s$ is the ratio of the total system bandwidth to the bandwidth required for an individual user and where N is the minimum integer value satisfying (15.8).

Example 15.2: Consider a TDMA cellular system with hexagonally shaped cells and with path-loss exponent $\gamma = 2$ for all signal propagation in the system. Find the minimum reuse factor N needed for a target SIR of 10 dB, and find the corresponding user capacity assuming a total system bandwidth of 20 MHz and a required signal bandwidth of 100 kHz.

Solution: To obtain the reuse factor, we apply (15.8) with $a_1 = .167$ and $a_2 = 3$ to obtain

$$N \geq \frac{\text{SIR}_0}{a_1 a_2} = \frac{10}{.5} = 20. \tag{15.10}$$

Now setting $G = B/B_s = 20 \cdot 10^6/100 \cdot 10^3 = 200$, we get $C_u = G/N = 10$ users per cell that can be accommodated. Typically $\gamma > 2$, as we consider in the next example.

Example 15.3: Consider a TDMA cellular system with diamond-shaped cells, pathloss exponent $\gamma = 4$ for all

signal propagation in the system, and BPSK modulation. Assume that the received signal exhibits Rayleigh fading. Suppose the users require $\bar{P}_b = 10^{-3}$. Assuming the system is interference limited, find the minimum reuse factor N that will meet this performance requirement. Also find the user capacity assuming a total system bandwidth of 20 MHz and a per-user signal bandwidth of 100 kHz.

Solution: Treating interference as Gaussian noise, in Rayleigh fading we have $\bar{P}_b \approx .25/\mathrm{SIR}_0$ for SIR_0 the average SIR ratio. The SIR required to meet the $\bar{P}_b$ target is thus $\mathrm{SIR}_0 = .25/10^{-3} = 250$ (approximately 24 dB). Substituting $\mathrm{SIR}_0 = 250$, $a_1 = .125,\ a_2 = 4$, and $\gamma = 4$ into (15.8) yields

$$N \geq \frac{1}{4}\sqrt{\frac{250}{.125}} = 11.18. \tag{15.11}$$

So a reuse factor of $N = 12$ meets the performance requirement. For the user capacity we have $G = B/B_s = 200$, so $C_u = G/N = 16$ users per cell can be accommodated. Note that the Gaussian assumption for the interference is just an approximation, which (by the central limit theorem) becomes more accurate as the number of interferers grows.

15.2.2 Nonorthogonal Systems (CDMA)

In nonorthogonal systems, codes (i.e. channels) are typically reused in every cell, so the reuse factor is $N = 1$. Since these systems exhibit both intercell and intracell interference, the user capacity is determined by the maximum number of users per cell that can be accommodated for a given target SIR. We will neglect intercell interference from outside the first tier of interfering cells – that is, from cells that are not adjacent to the cell of interest. We will also assume that all signals follow the simplified path-loss model with the same path-loss exponent. This assumption is typically true for interference from adjacent cells, but it ultimately depends on the propagation environment.

Let $N_c = N_T = C_u$ denote the number of channels per cell. In CDMA systems the user capacity is typically limited by the uplink, owing to the near–far problem and the asynchronicity of the codes. Focusing on the uplink, under the simplified path-loss model the received signal power is $P_r = P_t d^{-\gamma}$, where d is the distance between the mobile and its base station. There are $N_c - 1$ asynchronous intracell interfering signals and MN_c asynchronous intercell interfering signals transmitted from mobiles in the M adjacent cells. Let d_i, $i = 1, \ldots, N_c - 1$, denote the distance from the ith intracell interfering mobile to the uplink receiver and let P_i denote this mobile's transmit power. Let d_j, $j = 1, \ldots, MN_c$, denote the distance from the jth intercell interfering mobile to the uplink receiver and let P_j denote this mobile's transmit power. By Section 13.4.3, all interference is reduced by the spreading code cross-correlation $\xi/3G$, where G is the processing gain of the system and ξ is a parameter of the spreading codes with $1 \leq \xi \leq 3$. The total intracell and intercell interference power is thus given by

$$I = \frac{\xi}{3G}\left(\sum_{i=1}^{N_c-1} P_i d_i^{-\gamma} + \sum_{j=1}^{MN_c} P_j d_j^{-\gamma}\right), \tag{15.12}$$

which yields the ratio

$$\mathrm{SIR} = \frac{P_t d^{-\gamma}}{\frac{\xi}{3G}\left(\sum_{i=1}^{N_c-1} P_i d_i^{-\gamma} + \sum_{j=1}^{MN_c} P_j d_j^{-\gamma}\right)}. \tag{15.13}$$

Because all the distances in (15.13) are different, the expression cannot in general be further simplified without additional assumptions. Let us therefore assume perfect power control within a cell, so that the received power of

the desired signal and interfering signals within a cell are the same: $P_r = P_t d^{-\gamma} = P_i d_i^{-\gamma}$ for all i. Furthermore, let

$$\lambda = \frac{\sum_{j=1}^{MN_c} P_j d_j^{-\gamma}}{(N_c - 1)P_r} \tag{15.14}$$

denote the ratio of average received power from all intercell interference to that of all intracell interference under this power control algorithm. Using these approximations, we get the following formula for SIR, which is commonly used for the uplink SIR in CDMA systems with power control [10, 11]:

$$\text{SIR} = \frac{1}{\frac{\xi}{3G}(N_c - 1)(1 + \lambda)}. \tag{15.15}$$

Under this approximation, for any SIR target SIR_0 we can determine the user capacity $C_u = N_c$ by setting (15.15) equal to the target SIR and solving for C_u, which yields

$$C_u = 1 + \frac{1}{\frac{\xi}{3G}(1 + \lambda)\text{SIR}_0}. \tag{15.16}$$

Voice signals, because of their statistical nature, need not be continuously active [12]. The fraction of time that a voice user actually occupies the channel is called the *voice activity factor* and is denoted by α, where $0 < \alpha \leq 1$. If the transmitter shuts off during nonactivity then the interference in CDMA – that is, the denominator of (15.16) – is multiplied by α. This increases SIR and therefore user capacity.

Example 15.4: Consider a CDMA cellular system with perfect power control within a cell. Assume a target SIR_0 of 10 dB, a processing gain $G = 200$, spreading codes with $\xi = 2$, and equal average power from inside and outside the cell ($\lambda = 1$). Find the user capacity of this system.
Solution: From (15.16) we have

$$C_u = 1 + \frac{1}{\frac{2}{600}(2 \cdot 10)} = 16, \tag{15.17}$$

so 16 users per cell can be accommodated.

Since (15.16) and (15.9) provide simple expressions for user capacity, it is tempting to compare them for a given SIR target to determine whether TDMA or CDMA can support more users per cell. This was done in Examples 15.2 and 15.4, where for the same SIR target and other system parameters TDMA yielded 10 channels per cell while CDMA yielded 16. However, these capacity expressions are extremely sensitive to the modeling and system assumptions. Increasing λ from 1 to 2 in Example 15.4 reduces C_u for CDMA from 16 to 11, and changing the path-loss exponent in Example 15.2 for TDMA from $\gamma = 2$ to $\gamma = 3$ changes the reuse factor N from 20 to 6, which in turn changes user capacity from 10 to 33. Code-division multiple access systems can trade off spreading and coding, which yields high coding gains and a resulting lower SIR target at the expense of some processing gain [13]; high coding gain is harder to achieve in a TDMA system, since it cannot be traded for spreading gain. Voice activity was not taken into account for the CDMA system, which would lead to a higher capacity. Moreover, the CDMA capacity is derived under an assumption of perfect power control via channel inversion, whereas no power control is assumed for TDMA. The effects of shadowing and fading are also not taken into account, but fading will cause a power penalty in CDMA (owing to the channel inversion power control) and will also affect the intercell interference power for both TDMA and CDMA. All of these factors and trade-offs significantly complicate the analysis, which makes it difficult to draw general conclusions about the superiority of one technique over another

in terms of user capacity. Analysis of user capacity for both TDMA and CDMA under various assumptions and models associated with operational systems can be found in [11, 14, 15, 16].

15.3 Dynamic Resource Allocation

Cellular systems exhibit significant dynamics in the number of users in any given cell and in their time-varying channel gains. Moreover, users have different data-rate requirements that vary with time. Thus, resource allocation to users has evolved from the fixed resource allocation schemes in early cellular system to flexible methods that can support heterogeneous applications. In dynamic resource allocation the channels and power levels in the system are dynamically assigned relative to the current system conditions and user needs. Much work has gone into investigating dynamic resource allocation in cellular systems. In this section we summarize some of the main techniques, including dynamic channel allocation and power control.

15.3.1 Dynamic Channel Assignment

Dynamic channel assignment (DCA) falls into two categories: dynamic assignment of channels within a cell (intracell DCA) and, for orthogonally channelized systems, dynamic assignment of channels between cells (intercell DCA). Intercell DCA is typically not applicable to CDMA systems because channels are reused in every cell. The basic premise of intercell DCA is to make every channel available in every cell, so no fixed channel reuse pattern exists. Each channel can be used in every cell as long as the SIR requirements of each user are met. Thus, channels are assigned to users as needed, and when a call terminates the channel is returned to the pool of available channels for assignment. Intercell DCA has been shown to improve channel reuse efficiency by a factor of 2 or more over fixed reuse patterns, even with relatively simple algorithms [57, 58]. Mathematically, intercell DCA is a combinatorial optimization problem, with channel reuse constraints based on the SIR requirements. Most dynamic channel allocation schemes assume that all system parameters are fixed except for the arrival and departure of calls [59, 60, 61], with channel reuse constraints defined by a connected graph that is constant over time. The channel allocation problem under this formulation is a generalization of the vertex coloring problem and is thus NP-hard [62]. Reduced complexity has been obtained by applying neural networks [63] and simulated annealing [64] to the problem. However, these approaches can suffer from lack of convergence. The superior efficiency of DCA between cells is most pronounced under light loading conditions [59]. As traffic becomes heavier, DCA between cells can suffer from suboptimal allocations that are difficult to reallocate under heavy loading conditions. User mobility also affects performance of intercell DCA, since it causes more frequent channel reassignments and a corresponding increase in dropped calls [65]. Finally, the complexity of DCA between cells, particularly in heterogeneous networks with both large and small cells and when there are rapidly changing propagation conditions and user demands, limits the practicality of such techniques.

Intracell DCA allows dynamic assignment of channels within a cell to different users through scheduling, as discussed in Section 14.4. Schedulers must be efficient in their allocation of resources but also fair. There is generally a trade-off between these two goals since, as discussed in Section 14.4, the most efficient allocation of resources exploits multiuser diversity to allocate resources to the user with the best channel. However, this assumes that the users with the best channels can fully utilize these resources, which may not be the case if their application does not require a high data rate. Moreover, such an allocation is unfair to users with inferior channels. The same scheduling algorithms discussed in Section 14.4 - round robin, throughput maximization via MaxRate scheduling, and proportional fair scheduling - are commonly used in cellular systems. Performance of these schemes has been analyzed for a TDMA downlink in [45, 46]. Scheduling issues for the CDMA downlink and uplink –including rate adaptation, fairness, and deadline constraints- have been explored in [47, 48, 49] and [50, 51, 52], respectively. MIMO provides another degree of freedom in scheduling system resources, as outlined in [54, 55]. Multiple

classes of users can also be supported via appropriate scheduling, as described in [56]. An overview of the vast literature on scheduling algorithms for OFDMA cellular networks with MIMO and different user classes can be found in [53]. The specific choice of scheduling algorithm depends on the service provider's goals and tradeoffs for allocating their available spectrum. Service providers typically develop their own variations of the standard scheduling algorithms depending on their revenue model and customer agreements, including different classes of service.

15.3.2 Power Control

Power control is a critical aspect of wireless system design. As seen in prior chapters, a water-filling power adaptation maximizes capacity of adaptive-rate single-user and multiuser systems in fading. Power control via channel inversion maintains a fixed received SNR in single-user fading channels and also eliminates the near-far effect in CDMA uplinks. However, in cellular systems these power control policies affect intercell interference in different ways, as we now describe.

Power control on the downlink has less impact on intercell interference than on the uplink, since the downlink transmissions all originate from the cell center whereas uplink transmissions can come from the cell boundaries, which exacerbates interference to neighboring cells. Thus we will focus on the effect of power control on the uplink. Consider the two cells shown in Figure 15.8. Suppose that both mobiles B_1 and B_2 in cell B transmit at the same power. Then the interference caused by the mobile B_1 to the base station in cell A will be relatively large, since it is close to the boundary of cell A, while the interference from B_2 will generally be much weaker owing to the longer propagation distance. If water-filling power adaptation is employed, then B_1 will generally transmit at a lower power than B_2 because it will typically have a worse channel gain than B_2 to the base station in cell B, since it is farther away. This has the positive effect of reducing the intercell interference to cell A. In other words, water-filling power adaptation reduces intercell interference from mobiles near cell boundaries, the primary source of this interference. A similar phenomenon happens with multiuser diversity, since users transmit only when they have a high channel gain to their base station, which is generally true when they are close to their cell center. Conversely, under channel inversion the boundary mobiles will transmit at a higher power in order to maintain the same received power at the base station as mobiles that are near the cell center. This has the effect of increasing intercell interference from boundary mobiles.

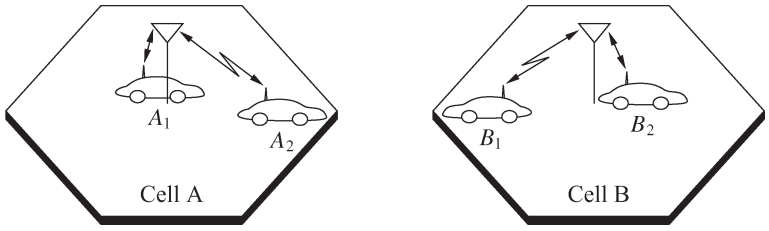


Figure 15.8: Effect of power control on intercell interference.

The power control algorithm discussed in Section 14.5 for maintaining a given SINR target for all users within a single cell can also be extended to multiple cells. Assume there are K users in the system with SINR requirement γ_k^* for the kth user. Focusing on the uplink, the kth user's SIR is given by

$$\gamma_k = \frac{g_k P_k}{n_k + \rho \sum_{k \neq j} g_{kj} P_j}, \; k, j \in \{1, \ldots, K\}, \tag{15.18}$$

where g_k is the channel power gain from user k to his base station, $g_{kj} > 0$ is the channel power gain from the jth (intercell or intracell) interfering transmitter to user k's base station, P_k is user k's transmit power, P_j is the

jth interferer's transmit power, n_k is the noise power at user k's base station, and ρ is the interference reduction due to signal processing (i. e., $\rho \approx 1/G$ for CDMA and $\rho = 1$ for TDMA) . Similar to the case of uplink power control, the SIR constraints can be represented in matrix form as $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ with $\mathbf{P} > 0$, where $\mathbf{P} = (P_1,\ P_2,\ \ldots,\ P_K)^T$ is the column vector of transmit powers for each user,

$$\mathbf{u} = \left(\frac{\gamma_1^* n_1}{g_1},\ \frac{\gamma_2^* n_2}{g_2},\ \ldots,\ \frac{\gamma_K^* n_K}{g_K} \right)^T \tag{15.19}$$

is the vector of noise powers scaled by the SIR constraints and channel gain, and $\mathbf{F}$ is an irreducible matrix with nonnegative elements given by

$$F_{kj} = \begin{cases} 0 & k = j, \\ \gamma_k^* g_{kj} \rho / g_k & k \neq j, \end{cases} \tag{15.20}$$

with $k,\ j \in \{1, 2,\ \ldots,\ K\}$. As in uplink power control for a single cell (see Section 14.5), it is shown in [68, 69, 70] that if the Perron-Frobenius (maximum modulus) eigenvalue of $\mathbf{F}$ is less than unity then there exists a vector $\mathbf{P} > 0$ (i.e., $P_k > 0$ for all k) such that the SIR requirements of all users in all cells are satisfied, with $\mathbf{p}^* = (\mathbf{I} - \mathbf{F})^{-1}\mathbf{u}$ the Pareto optimal solution. Thus, $\mathbf{p}^*$ meets the SIR requirements with the minimum transmit power of the users. Moreover, the distributed iterative power control algorithm,

$$P_k(i+1) = \frac{\gamma_k^*}{\gamma_k(i)} P_k(i), \tag{15.21}$$

converges to the optimal solution. This is a simple algorithm for power control because it requires only SIR information at each transmitter, where each transmitter increases power when its SIR is below its target and decreases power when its SIR exceeds its target. However, it is important to note that the existence of a feasible power control vector that meets all SIR requirements is less likely in a cellular system than in the single-cell uplink, since there are more interferers contributing to the SIR and since the channel gains range over a larger set of values than in a single-cell uplink. When there is no feasible power allocation, the distributed algorithm will result in all users transmitting at their maximum power and still failing to meet their SIR requirement.

Power control is often combined with scheduling, intercell DCA, or intracell DCA. Intercell DCA and power control for TDMA are analyzed in [71, 72], and these techniques are extended to MIMO systems in [73]. Power control combined with intracell DCA exploiting multiuser diversity is investigated in [74, 75]. Intracell DCA for CDMA via variable spreading is combined with power control in [76, 77]. A comparison of power control combined with either multicode or variable-spreading CDMA is given in [78].

15.4 Advanced Network Architectures

15.4.1 Heterogeneous Networks (Hetnets)

Early cellular networks were designed with only large cells, also called macrocells, which are on the order of a kilometer within a city, several kilometers in a suburban area, and tens of kilometers in a rural area. These macrocells are created by placing base stations on top of buildings, large towers, or mountains. The base stations and mobiles within a macrocell must transmit at a sufficiently high power to maintain the SINR level needed for voice or data communication. Cellular networks with macrocells can cover an entire city with a relatively small number of base stations, which minimizes deployment and upgrade time and cost. One disadvantage of having only macrocells in a cellular network is the high transmit power required to achieve the desired SINR, which reduces mobile battery life and increases base station power cost. In addition, cellular networks with only macrocells have lower capacity than those with smaller cells, as discussed in Section 15.1.3. Finally, fading and shadowing can lead

to coverage holes (outages) in a macrocell, which may be filled by increasing the transmit power at the base station or mobile experiencing the outage at the expense of increased interference to co-channel users. If the transmission path between the mobile and its base station is completely blocked, outage at the mobile cannot be avoided.

These shortcomings in macrocell cellular network architectures can be addressed through a hybrid approach utilizing both small and large cells. This heterogenous network, or Hetnet, deploys an underlay of small cells beneath the macrocell network, with small cells placed in locations where more capacity, coverage, or power efficiency is needed, as shown in Fig. 15.9. The small cells, which can be indoor or outdoor, are typically deployed in areas with a high density of users such as downtown districts, sports arenas, airports, train stations, and shopping malls.

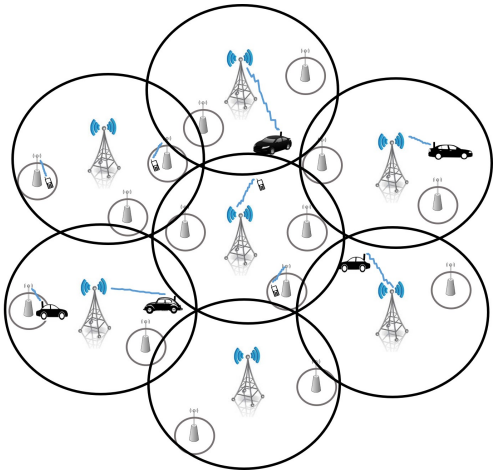


Figure 15.9: Hetnet Cellular Architecture with Macrocells and Small Cells

The design and performance analysis in a Hetnet must take into account the irregularly-placed small cells as well as the regularly-placed macrocells. From a design perspective, channel allocation to large and small cells along with power control must ensure the SINR requirements for each user are met. The simplest paradigm for channel allocation in Hetnets assigns different channel sets to the large and small cells so they don't interferer with each other. In this case, the fixed or dynamic macrocell channel allocation algorithms described, respectively, in Chapters 15.1.4 and 15.3.1 are used to allocate the assigned channel set across the large cells. The sparse and irregular placement of small cells precludes the use the fixed channel allocation method of Chapter 15.1.4. Instead, fixed channel allocation across small cells is done based on SIR field measurements at the time of deployment or, alternatively, through dynamic channel allocation based on real-time SIR receiver estimates. The sharing of some or all channels across the large and small cells increases the complexity of the channel allocation problem. Often small cells are assigned higher frequency channels, e.g. in the millimeter wave band, since power falls off quickly relative to distance in such channels. This leads to less interference between small cells utilizing the same high-frequency channel.

The large or small cell base station that provides a user with the highest SINR may not be the one associated with the highest received signal power as that base station-user channel may also exhibit high interference. Thus, dynamic resource allocation in Hetnets typically optimizes which large or small cell base station a user should connect to along with the associated channel and transmit power. Resource allocation in Hetnets is highly complex given the large number of parameters involved as well as site-specific topologies and user dynamics. To tame the complexity, resource allocation algorithms utilize advanced optimization methods [79], game theory [80], or machine learning [81]. A cellular network that autonomously configures and dynamically optimizes its parameters across all of its base stations and users, including channels, power levels, and user-base station associations, is

referred to as a self-organizing network (SON) [82].

Once resources have been allocated within the Hetnet, determining the SIR or SINR and the associated performance for any given user is more challenging than in a system with homogeneous cell sizes. Simple analytical models are not used for this analysis as they cannot easily capture the effect of irregularly-placed small cells. As with macrocell networks, extensive measurements of deployed systems as well as simulations that average over variations in the large and small base station and user locations as well as the user and interferer channels are typically used to determine system performance [23, 24]. The stochastic geometry analytical approach for determining SINR can also be extended to Hetnet cellular architectures [25, 26]. In [24] this analytical approach is validated by simulations, where it is demonstrated that deploying small cells within a macrocell network increases user average throughput by approximately a factor of 4 even without interference coordination.

15.4.2 Cloud Radio Access Networks (C-RAN)

The basic premise of a Cloud Radio Access Network (C-RAN) is to design base stations to consist of just radios and analog signal processing, with baseband signal processing done in the cloud. Then, rather than have each baseband processing unit (BBU) connected to just one base station remote radio head (RRH), the BBUs sit within the network cloud so their resources can be pooled and allocated dynamically. The BBU pool generates a baseband signal that is distributed to one or more RRHs for analog signal processing and transmission through the base station antenna(s). The RRHs are connected via a reliable high-speed low-latency link to the BBU cloud so as to replicate co-location of the analog and digital signal processing that base stations perform in conventional cellular architectures. This C-RAN architecture is illustrated in Fig. 15.10.

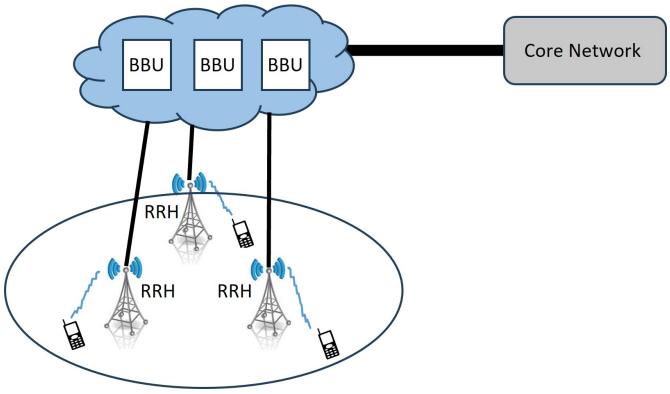


Figure 15.10: Cloud RAN Cellular Network Architecture

The BBU pool performs the baseband signal processing for the digital modulation and coding associated with data transmission and reception. Each RRH operates on the baseband modulated digital signal generated by one or more BBUs. In particular, for the downlink (base station to user device), the BBU pool generates a baseband modulated digital signal which is sent to one or more RRHs for transmission. The RRH performs D/A conversion on the baseband modulated digital signal, upconverts the resulting analog baseband signal to the carrier frequency, amplifies it, and then transmits the resulting analog signal through its antenna. For the uplink (user device to base station), the RRH receives the analog signal through its antenna, downconverts it to baseband, and then performs A/D conversion to obtain the received baseband modulated digital signal which is then transmitted to the BBU pool. The BBU performs the baseband processing functions, including modulation and demodulation as well as coding and decoding.

The main benefit of the C-RAN architecture is its flexible allocation of resources, as the BBUs can be dynamically allocated to different RRHs as network conditions and user requirements change. Significant performance

gains in spectrum and energy efficiency relative to conventional cellular architectures can be obtained by exploiting this flexibility [27, 28, 29, 30]. In addition, cooperation across base stations in terms of handoff and joint transmission are easily implemented in the C-RAN architecture since the BBU pool supports such coordinated base station signal processing [31, 32]. The main disadvantage of the C-RAN architecture is the latency introduced by performing baseband signal processing in the cloud instead of at the base station, as is done in conventional cellular architectures. To address this issue, the C-RAN design typically imposes stringent latency and speed requirements on the links between the BBU pool and RRHs [33, 34]. If these link requirements are not met, the C-RAN network will perform poorly.

The C-RAN network architecture can be enhanced with network virtualization and open hardware. Wireless network virtualization consists of software in the cloud or core network that creates multiple virtual networks offering different user services that run over the same physical network infrastructure [35]. Most C-RAN architectures are deployed by a single operator using a single vendor for all of the network equipment (RRHs, BBUs, and all of the cloud software). A more flexible architecture is open RAN (O-RAN), which is a C-RAN architecture with open interfaces such that the RRHs and BBUs can come from different vendors [36, 37].

15.5 Fundamental Rate Limits

15.5.1 Shannon Capacity of Cellular Systems

There have been few information-theoretic results on the Shannon capacity of cellular systems, which is due to the difficulty of incorporating channel reuse and the resulting interference into fundamental capacity analysis. Whereas the capacity for the uplink and downlink of an isolated cell is known, as described in Sections 14.6 and 14.7, there is little work on extending these results to multiple cells. The capacity has been characterized in some cases, but the capacity as well as optimal transmission and reception strategies under broad assumptions about channel modeling, base station cooperation, interference characteristics, and transmitter/receiver CSI remain mostly unsolved.

One way to analyze the capacity of a cellular system is to assume that the base stations fully cooperate to jointly encode and decode all signals. In this case the notion of cells does not enter into capacity analysis. Specifically, under the assumption of full base station cooperation, the multiple base stations can be viewed as a single base station with multiple geographically dispersed antennas. Transmission from the multiple-antenna base station to the mobiles can be treated as a MIMO downlink (broadcast channel), and transmission from the mobiles to the multiple-antenna base station can be treated as a MIMO uplink (multiple access channel). The Shannon capacity regions for both of these channels are known, as discussed in Section 14.10, for some channel models and assumptions about channel side information.

The uplink capacity of cellular systems under the assumption of full base station cooperation was first investigated in [83], followed by a more comprehensive treatment in [84]. In both works, propagation between the mobiles and the base stations was characterized using an AWGN channel model, with a channel gain of unity within a cell and a gain of α, $0 \leq \alpha \leq 1$, between cells. The Wyner model of [84] considers one-and two-dimensional arrays of cells and in both cases derives the per-user capacity – defined as the maximum possible rate that all users can maintain simultaneously – as

$$C(\alpha) = \frac{B}{K} \int_0^1 \log_2 \left(1 + \frac{KP(1 + 2\alpha \cos(2\pi\theta))^2}{N_0 B} \right) d\theta, \tag{15.22}$$

where B is the total system bandwidth, $N_0/2$ is the noise power spectral density, K is the number of mobiles per cell, and P is the average transmit power of each mobile. It is also shown in both [84] and [83] that uplink capacity is achieved by using orthogonal multiple access techniques (e.g. TDMA) in each cell and then reusing these orthogonal channels in other cells, although this is not necessarily uniquely optimal. The behavior of $C(\alpha)$

as a function of α, the attenuation of the intercell interference, depends on the SNR of the system. The per-user capacity $C(\alpha)$ generally increases with α at high SNRs, since having strong intercell interference aids in decoding and subsequent subtraction of the interference from desired signals. However, at low SNR, $C(\alpha)$ initially decreases with α and then increases. This is because weak intercell interference cannot be reliably decoded and subtracted, so such interference reduces capacity. As the channel gains associated with the intercell interference grows, the joint decoding is better able to decode and subtract out this interference, leading to higher capacity.

An alternate analysis method for capacity of cellular systems assumes no base station cooperation, so that the receivers in each cell treat signals from other cells as interference. This approach mirrors the design of cellular systems in practice. Unfortunately, Shannon capacity of channels with interference is a long-standing open problem in information theory [85, 86], solved only for the special case of strong interference [87]. By treating the interference as Gaussian noise, the capacity of both the uplink and downlink can be determined using the single-cell analysis of Sections 14.6 and 14.7. The Gaussian assumption can be viewed as a worst-case assumption about the interference, since exploiting any known structure of the interference can presumably help in decoding the desired signals and therefore increases capacity. The capacity of a cellular system uplink with fading based on treating interference as Gaussian noise was obtained in [88] for both one- and two-dimensional cellular grids. These capacity results show that, with or without fading, if intercell interference is nonnegligible then an orthogonal multiple access method (e.g. TDMA) within a cell is optimal. This generalization holds also when channel-inversion power control is used within a cell. Moreover, in some cases limited or no reuse of channels in different cells can increase capacity. The effects on capacity for this model when there is partial joint processing between base stations have also been characterized [89].

The results described here provide some insight into the capacity and optimal transmission strategies for the uplink of cellular systems. Unfortunately, no such results are yet available for the downlink under any assumptions about channel modeling or base station cooperation. Although the uplink was the capacity bottleneck for cellular systems providing two-way voice, the downlink is becoming increasingly critical for multimedia downloads. Therefore, a better understanding of the capacity limitations and insights for cellular downlinks would be most beneficial in future cellular system design.

15.5.2 Area Spectral Efficiency

The Shannon capacity regions for cellular systems described in Section 15.5.1 dictate the set of maximum achievable rates on cell uplinks or downlinks. When the capacity region is computed based on the notion of joint processing at the base stations, then there is effectively only one cell with a multiple-antenna base station. However, when capacity is computed based on treating intercell interference as Gaussian noise, the capacity region of both the uplink and downlink becomes highly dependent on the cellular system structure – in particular, the cell size and channel reuse distance. Area spectral efficiency (ASE) is a capacity measure that allows the cellular structure (and in particular the reuse distance) to be optimized relative to fundamental capacity limits.

Recall that, for both orthogonal and nonorthogonal channelization techniques, the reuse distance D in a cellular system defines the distance between any two cell centers that use the same channel. Since these channels are reused at distance D, the area covered by each channel is roughly the area of a circle with radius $.5D$, which is given by $A = \pi(.5D)^2$. The larger the reuse distance, the less efficient the channel reuse. However, reducing the reuse distance increases intercell interference, thereby reducing the capacity region of each cell if this interference is treated as noise. The ASE captures this trade-off between efficient resource use and the capacity region per cell.

Consider a cellular system with K users per cell, a reuse distance D, and a total bandwidth allocation B. Let $\mathcal{C} = (R_1, R_2, \ldots, R_K)$ denote the capacity region, for either the uplink or the downlink, in a given cell when the intercell interference from other cells is treated as Gaussian noise. The corresponding sum-rate, also called the system throughput, is given by

$$\mathcal{C}_{\rm SR} = \max_{(R_1,\ldots,R_K)\in\mathcal{C}} \sum_{k=1}^{K} R_k \text{ bps}. \tag{15.23}$$

The region $\mathcal{C}$ and corresponding sum-rate $\mathcal{C}_{\rm SR}$ can be obtained for any channelization technique within a cell. Clearly, this capacity region will decrease as intercell interference increases. Moreover, since intercell interference decreases as the reuse distance increases, the size of the capacity region will increase with reuse distance.

The ASE of a cell is defined as the throughput per hertz per unit area that is supported by a cell's resources. Specifically, given the sum-rate capacity just described, the ASE is defined as

$$A_e \triangleq \frac{\mathcal{C}_{\rm SR}/B}{\pi(.5D)^2} \text{ bps/Hz/m}^2. \tag{15.24}$$

From [88], orthogonal channelization is capacity achieving within a cell, so we will focus on TDMA for computing ASE. If we also assume that the system is interference limited, so that noise can be neglected, then the rate R_k associated with each user in a cell is a function of his received signal-to-interference power $\gamma_k = P_k/I_k$, $k = 1,\ldots,K$. If γ_k is constant then $R_k = \tau_k B \log(1+\gamma_k)$, where τ_k is the time fraction assigned to user k. Typically γ_k is not constant, since both the interference and signal power of the kth user will vary with propagation conditions and mobile locations. When γ_k varies with time, the capacity region is obtained from optimal resource allocation over time and across users, as described in Sections 14.6.4 and 14.7.2.

As a simple example, consider an AWGN TDMA uplink with hexagonal cells of radius R. Assume all users in the cell are assigned the same fraction of time $\tau_k = 1/K$ and have the same transmit power P. We neglect the impact of intercell interference outside the first tier of interfering cells, so there are six intercell interferers. We take a pessimistic model, where (i) all users in the cell of interest are located at the cell boundary and (ii) all intercell interferers are located at their cell boundaries closest to the cell of interest. Path loss is characterized by the simplified model $P_r = P_t d^{-2}$ within a cell and by $P_r = P_t d^{-\gamma}$ between cells, where $2 \le \gamma \le 4$. The received signal power of the kth user is then $P_k = PR^{-2}$, and the intercell interference power is $I_k = 6P(D-R)^{-\gamma}$. The maximum achievable rate for the kth user in the cell is thus

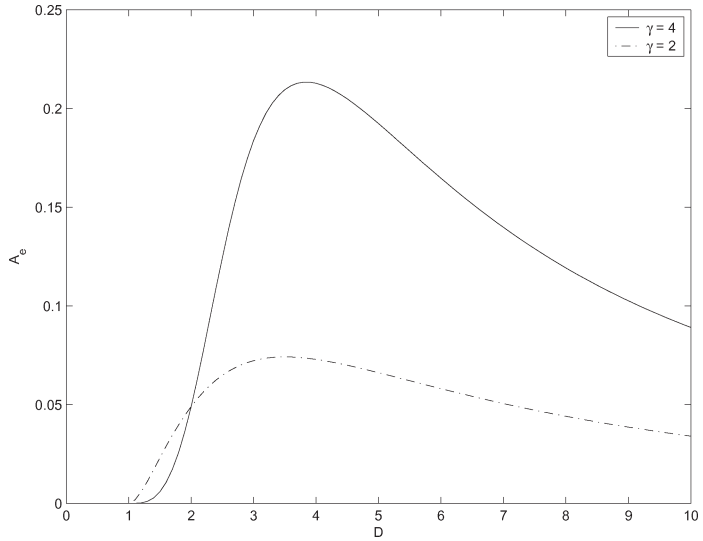


Figure 15.11: Area spectral efficiency for AWGN uplink ($\gamma = 2$ and 4).

$$R_k = \frac{B}{K} \log_2 \left(1 + \frac{(D-R)^\gamma}{6R^2} \right) \text{ bps}, \tag{15.25}$$

and the ASE is

$$A_e = \frac{\sum_{k=1}^{K} R_k/B}{\pi(.5D)^2} = \frac{\log_2(1 + (D-R)^\gamma/6R^2)}{\pi(.5D)^2} \text{ bps/Hz/m}^2. \tag{15.26}$$

Plots of A_e versus D for $\gamma = 4$ and $\gamma = 2$ are shown in Figure 15.11, with the cell radius normalized to $R = 1$. Comparing these plots we see that, as expected, if the intercell interference path loss falls off more slowly then the ASE is decreased. However, it is somewhat surprising that the optimal reuse distance is also decreased.

Suppose now that the interferers are not on the cell boundaries. If all interferers are at a distance $D - R/2$ from the desired user's base station, then the ASE becomes

$$A_e = \frac{\log_2(1 + (D-R/2)^\gamma/6R^2)}{\pi(.5D)^2}. \tag{15.27}$$

The ASE in this case is plotted in Figure 15.12 for $\gamma = 4$, along with the ASE for interferers on their cell boundaries. As expected, the ASE is larger for interferers closer to their cell centers than on the boundaries closest to the cell they interfere with, and the optimal reuse distance is smaller.

Area spectral efficiency has been characterized in [90] for a cellular system uplink with orthogonal channelization assuming variable-rate transmission as well as best-case, worst-case, and average intercell interference conditions. The impact of different fading models, cell sizes, and system load conditions was also investigated. The results indicate that the optimal reuse factor is unity for both best-case and average interference conditions. That is, channels should be reused in every cell even though there is no interference reduction from spreading. Moreover, the ASE decreases as an exponential of a fourth-order polynomial relative to the cell radius, thus quantifying the capacity gains associated with reducing cell size. A similar framework was used in [91] to characterize the ASE of cellular downlinks.

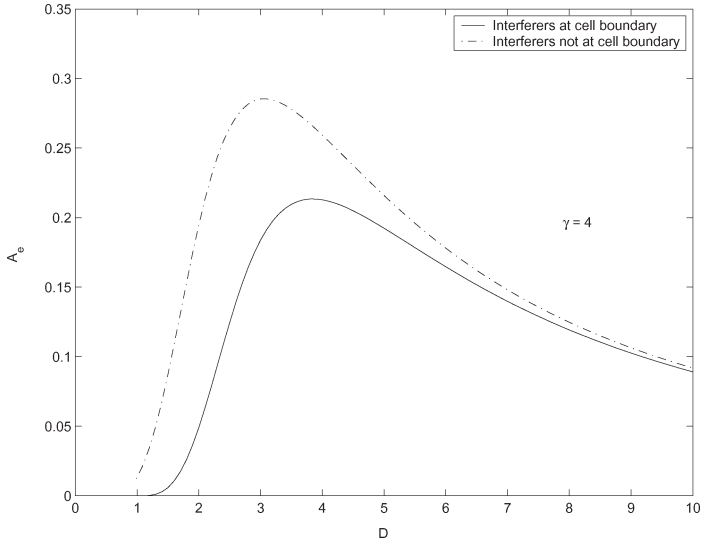


Figure 15.12: ASE for interferers at distance $D - R/2$ and at distance D ($\gamma = 4$).

Chapter 15 Problems

1. Consider a city of 10 square kilometers. A macrocellular system design divides the city into square cells of 1 square kilometer, where each cell can accommodate 100 users. Find the total number of users that can be accommodated in the system and the length of time it takes a mobile user to traverse a cell (approximate time needed for a handoff) when moving at 30 km per hour. If the cell size is reduced to 100 square meters and everything in the system scales so that 100 users can be accommodated in these smaller cells, find the total number of users the system can accommodate and the length of time it takes to traverse a cell.

2. Show that the reuse distance $D = \sqrt{3}R\sqrt{i^2 + j^2 + ij}$ for the channel assignment algorithm associated with hexagonal cells that is described in Section 15.1.4.

3. Consider a cellular system with diamond-shaped cells of radius $R = 100\,\mathrm{m}$. Suppose the minimum distance between cell centers using the same frequency must be $D = 600\,\mathrm{m}$ to maintain the required SIR.

   (a) Find the required reuse factor N and the number of cells per cluster.

   (b) If the total number of channels for the system is 450, find the number of channels that can be assigned to each cell.

   (c) Sketch two adjacent cell clusters and show a channel assignment for the two clusters with the required reuse distance.

4. Consider a cellular system with hexagonal cells of radius $R = 1\,\mathrm{km}$. Suppose the minimum distance between cell centers using the same frequency must be $D = 6\,\mathrm{km}$ to maintain the required SIR.

   (a) Find the required reuse factor N and the number of cells per cluster.

   (b) If the total number of channels for the system is 1200, find the number of channels that can be assigned to each cell.

   (c) Sketch two adjacent cell clusters and show a channel assignment for the two clusters with the required reuse distance.

5. Compute the SIR for a TDMA cellular system with diamond-shaped cells, where the cell radius $R = 10\,\mathrm{m}$ and the reuse distance $D = 60\,\mathrm{m}$. Assume that the path-loss exponent within the cell is $\gamma_I = 2$ but that the intercell interference has path-loss exponent $\gamma_0 = 4$. Compare with the SIR for $\gamma = \gamma_I = \gamma_0 = 4$ and for $\gamma = \gamma_I = \gamma_0 = 2$. Explain the relative orderings of SIR in each case.

6. Find the minimum reuse distance and user capacity for a TDMA cellular system with hexagonally shaped cells, path-loss exponent $\gamma = 2$ for all signal propagation in the system, and BPSK modulation. Assume an AWGN channel model with required $P_b = 10^{-6}$, a total system bandwidth of $B = 50$ MHz, and a required signal bandwidth of 100 kHz for each user.

7. Consider a CDMA system with perfect power control, a processing gain of $G = 100$, spreading codes with $\xi = 1$, and the ratio $\lambda = 1.5$. For BPSK modulation with a target P_b of 10^{-6}, under an AWGN channel model find the user capacity of this system with no sectorization and with $N = 3$ sectors.

8. In this problem we consider the impact of voice activity factor, which creates a random amount of interference. Consider a CDMA system with

$$\mathrm{SIR} = \frac{G}{\sum_{i=1}^{N_c - 1} \chi_i + N};$$

here G is the processing gain of the system, the χ_i represent intracell interference and follow a Bernoulli distribution with probability $\alpha = p(\chi_i = 1)$ equal to the voice activity factor, and N characterizes intercell interference and is assumed to be Gaussian, with mean $.247N_c$ and variance $.078N_c$. The probability of outage is defined as the probability that the SIR is below some target SIR_0:

$$P_{\text{out}} = p(\text{SIR} < \text{SIR}_0).$$

(a) Show that

$$P_{\text{out}} = p\left(\sum_{i=1}^{N_c-1} \chi_i + N > \frac{G}{\text{SIR}_0}\right).$$

(b) Find an analytical expression for P_{out}.

(c) Using the analytical expression obtained in part (b), compute the outage probability for $N_c = 35$ users, $\alpha = .5$, and a target SIR of $\mathrm{SIR}_0 = 5(7\,\mathrm{dB})$. Assume $G = 150$.

(d) Assume now that N_c is sufficiently large that the random variable $\sum_{i=1}^{N_c-1} \chi_i$ can be approximated as a Gaussian random variable. Under this approximation, find the distribution of $\sum_{i=1}^{N_c-1} \chi_i + N$ as a function of N_c as well as an analytical expression for outage probability based on this approximation.

(e) Compute the outage probability using this approximation for $N_c = 35$, $\alpha = .5, G = 150$, and a target SIR of $\mathrm{SIR}_0 = 5(7\,\mathrm{dB})$. Compare with your results in part (c).

9. Assume a cellular system with K users, and suppose that the minimum SIR for each user on the downlink is given as $\gamma_1^*, \ldots, \gamma_K^*$. Write down the conditions such that a power control vector exists that will satisfy these constraints.

10. Plot the ASE versus reuse distance $D(0 \leq D \leq 10)$ of a TDMA uplink with hexagonal cells of radius $R = 1$, assuming that all users in the cell are assigned the same fraction of time and the same transmit power and that all users are located on a cell boundary. Your computation should be based on a path loss of $\gamma = 2$ and interference from only the first ring of interfering cells, where the interferers have probability 1/5 of being in one of the following five locations: the cell boundary closest to the cell of interest; halfway between the base station and this closest cell boundary; in the cell center; the cell boundary farthest from the cell of interest; or halfway between the base station and this farthest cell boundary. Also plot the ASE and optimal reuse distance for the cases where, with probability 1, the interferers are on the (a) closest or (b) farthest cell boundary. Compare the differences in these plots.

11. Consider a one-dimensional cellular system deployed along a highway. The system has square cells of length $2R = 2$ km, as shown in Figure 15.13. This problem focuses on downlink transmission from base stations to mobiles. Assume that each cell has two mobiles located as shown in the figure, so that the mobiles in each cell have the exact same location relative to their respective base stations. Assume a total transmit power at each base station of $P_t = 5\,\mathrm{W}$, which is evenly divided between the two users in its cell. The total system bandwidth is 100 kHz, and the noise power spectral density at each receiver is $10^{-16}\,\mathrm{W/Hz}$. The propagation follows the model $P_r = P_t K(d_0/d)^3$, where $d_0 = 1\,\mathrm{m}$ and $K = 100$. All interference should be treated as AWGN, and interference outside the first ring of interfering cells can be neglected. The system uses a frequency-division strategy, with the bandwidth allocated to each base station evenly divided between the two users in its cell.

(a) For a reuse distance $D = 2$ (frequencies reused every other cell), what bandwidth is allocated to each user in the system?

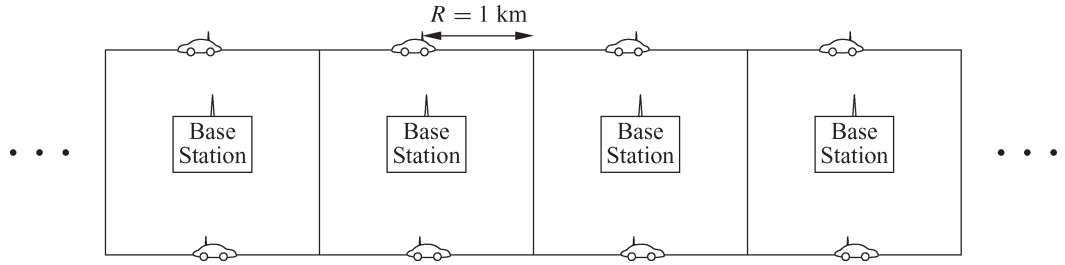


Figure 15.13: One-dimensional cellular system with square cells (Problem 15-11).

(b) Compute the minimum reuse distance D required to achieve a 10^{-3} BER for BPSK modulation in fast Rayleigh fading.

(c) Neglecting any fading or shadowing, use the Shannon capacity formula for user rates to compute the area spectral efficiency of each cell under frequency division, where frequencies are reused every other cell ($D = 2$).

12. In this problem we investigate the per-user capacity for the uplink of a cellular system for different system parameters. Assume a cellular system uplink where the total system bandwidth is $B = 100\,\text{kHz}$. Assume in each cell that the noise PSD $N_0/2$ at each base station receiver has $N_0 = 10^{-9}\,\text{W/Hz}$ and that all mobiles have the same transmit power P.

(a) Plot the per-user capacity for the uplink, as given by (15.22), for $K = 10$ users, transmit power $P = 10\,\text{mW}$ per user, and $0 \leq \alpha \leq 1$. Explain the shape of the curve relative to α.

(b) For $\alpha = .5$ and $P = 10\,\text{mW}$, plot the per-user capacity for $1 \leq K \leq 30$. Explain the shape of the curve relative to K.

(c) For $\alpha = .5$ and $K = 10$ users, plot the per-user capacity for $0 \leq P \leq 100\,\text{mW}$. Explain the shape of the curve relative to P.

13. As discussed in [24], stochastic analysis in HetNets is based on a Poisson spatial point process model for the base station locations in each tier of cells. In this model, the base stations in the ith tier are independently distributed throughout the possible spatial locations with density λi, which means that if the total area covered by the base stations in tier i is A, there are $\lambda i A$ base stations within this area. For tier i, the base station transmit power in tier i is denoted as P_i, the path loss exponent is α_i and the target SINR within the tier is $\$beta_i$. The most common Hetnet model - a macrocell network with a small cell underlay - consists of two tiers.

As derived in [25], the downlink outage probability for a mobile at the center of area A given a target SINR of β_i in tier i, assuming noise is negligible relative to interference, is $P_{out} = 1 - p[\text{SINR}_i > \beta_i]$ for

$$P_{out} = 1 - \frac{\pi}{C(\alpha)} \frac{\sum_{i=1}^{K} \lambda_i P_i^{2/\alpha} \beta_i^{-2/\alpha}}{\sum_{i=1}^{K} \lambda_i P_i^{2/\alpha}}, \tag{15.28}$$

where λ_i, P_i, α_i, and β_i are as described above and $C(\alpha) = (2\pi 2/\alpha) csc(2\pi/\alpha)$ is a simple constant. This derivation assumes that the mobile connects to the base station for which it has the strongest SNR, there is no interference management, each channel associated with tier i experiences path loss with exponent α_i along with Rayleigh fading, and the target SINR for each tier is at or above 0 dB. This does not necessarily result in maximizing the SINR at the mobile.

(a) Show that if the target SINRs for each tier are the same, i.e $\beta_i = \beta$, then this expression simplifies to

$$P_{out} = 1 - \frac{\pi}{C(\alpha)\beta^{2/\alpha}}. \tag{15.29}$$

Note that this simplified expression implies the surprising result that the downlink outage probability at the center of A is independent of the number and density of base stations in each tier, the number of tiers, the transmit power of base stations in each tier, and the fading distribution.

(b) For a two-tier network of macrocells ($i = 1$) and small cells ($i = 2$), suppose the path loss and Rayleigh fading propagation model is modified to include shadowing on the macrocell base stations only and all other model parameters are unchanged. Assuming mobiles still connect to the base station with the largest SNR, how will this change the outage probability at the center of area A (increase, decrease, or stay the same), and why?

(c) Under the same assumptions as part (b), how would adding one small cell in tier 2 affect outage probability at the center of area A (increase, decrease, or stay the same) and why?

Bibliography

[1] T. S. Rappaport, *Wireless Communications – Principles and Practice*, 2nd ed., Cambridge University Press, 2024.

[2] V. Erceg, A. J. Rustako, and R. S. Roman, "Diffiaction around corners and its effects on the microcell coverage area in urban and suburban environments at 900 MHz, 2 GHz, and 4 GHz," *IEEE Trans. Veh. Tech*., pp. 762–6, August 1994.

[3] A. J. Goldsmith and L. J. Greenstein, "A measurement-based model for predicting coverage areas of urban microcells," *IEEE J. Sel. Areas Commun*., pp. 1013–23, September 1993.

[4] S. Dehghan and R. Steele, "Small cell city," *IEEE Commun. Mag*., pp. 52–9, August 1997.

[5] H. Holma and A. Toskala, *WCDMA for UMTS Radio Access for Third Generation Mobile Communications*, 3rd ed., Wiley, New York, 2004.

[6] M. V. Clark, V. Erceg, and L. J. Greenstein, "Reuse efficiency in urban microcellular networks," *IEEE Trans. Veh. Tech*., pp. 279–88, May 1997.

[7] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*, Kluwer, Dordrecht, 2002.

[8] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed., Kluwer, Dordrecht, 2001.

[9] R. Prasad and A. Kegel, "Effects of Rician faded and log-normal shadowed signals on spectrum efficiency in microcellular radio," *IEEE Trans. Veh. Tech*., pp. 274–81, August 1993.

[10] R. Kohno, R. Meidan, and L. B. Milstein, "Spread spectrum access methods for wireless communications," *IEEE Commun. Mag*., pp. 58–67, January 1995.

[11] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech*., pp. 303–12, May 1991.

[12] P. T. Brady, "A statistical analysis of on–off patterns in 16 conversations," *Bell System Tech. J*., pp. 73–91, January 1968.

[13] V. V. Veeravalli and A. Mantravadi, "The coding–spreading trade-off in CDMA systems," *IEEE J. Sel. Areas Commun*., pp. 396–408, February 2002.

[14] P. Jung, P. W. Baier, and A. Steil, "Advantages of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications," *IEEE Trans. Veh. Tech*., pp. 357–64, August 1993.

[15] T. S. Rappaport and L. B. Milstein, "Effects of radio propagation path loss on DS-CDMA cellular frequency reuse efficiency for the reverse channel," *IEEE Trans. Veh. Tech*., pp. 231–42, August 1992.

[16] B. Gundmundson, J. Sköld, and J. K. Ugland, "A comparison of CDMA and TDMA systems," *Proc. IEEE Veh. Tech. Conf.*, pp. 732–5, May 1992.

[17] J. Winters, "Smart antennas for wireless systems," *IEEE Pers. Commun. Mag.*, pp. 23–7, February 1998.

[18] R. W. Heath, Jr., S. Peters, Y. Wang and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 161-167, April 2013.

[19] Elnashar and M. A. El-Saidny, "Looking at LTE in practice: A performance analysis of the LTE system based on field test results," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 81-92, Sept. 2013.

[20] K. I. Pedersen et al., "A tutorial on radio system-level simulations With emphasis on 3GPP 5G-advanced and beyond," *IEEE Communications Surveys and Tutorials*, vol. 26, no. 4, pp. 2290-2325, 2024.

[21] J.-P. M. G. Linnartz, "Exact analysis of the outage probability in multiple-user mobile radio," *IEEE Trans. Comm.*, vol. 40, no. 1, pp. 20-23, Jan. 1992.

[22] D. B. Cheikh, J.-M. Kelif, M. Coupechoux, and P. Godlewski, "SIR distribution analysis in cellular networks considering the joint impact of path-loss, shadowing and fast fading", *Eurasip J. Wireless Com. Network.* pp. 1-10, 2011.

[23] P. Ökvist and A. Simonsson, "LTE HetNet trial - range expansion including micro/pico indoor coverage survey," *Proc. 2012 IEEE Vehicular Technology Conference (VTC Fall)*, September 2012.

[24] A. Ghosh et al., "Heterogeneous cellular networks: From theory to practice," *IEEE Communications Magazine*, vol. 50, no. 6, pp. 54-64, June 2012,

[25] J.G. Andrews, A. K. Gupta, A. Alammouri, and H. S. Dhillon. *An Introduction to Cellular Network Analysis Using Stochastic Geometry*. Cham, Switzerland: Springer, 2023.

[26] B. Błaszczyszyn, M. Haenggi, P. Keeler, and S. Mukherjee, *Stochastic Geometry Analysis of Cellular Networks*, Cambridge University Press, 2018.

[27] T.Q.S Quek, M Peng, O. Simeone, W. Yu eds. *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge University Press; 2017.

[28] J. Wu, Z. Zhang, Y. Hong and Y. Wen, "Cloud radio access network (C-RAN): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35-41, Jan.-Feb. 2015.

[29] N. Moosavi, M. Sinaie, P. Azmi, P. -H. Lin and E. Jorswieck, "Cross Layer Resource Allocation in H-CRAN With Spectrum and Energy Cooperation," *IEEE Transactions on Mobile Computing*, vol. 22, no. 1, pp. 145-158, 1 Jan. 2023.

[30] W. Ejaz, S. K. Sharma, S. Saadat, M. Naeem, A. Anpalagan, and N.A. Chughtai. "A comprehensive survey on resource allocation for CRAN in 5G and beyond networks," *Journal of Network and Computer Applications*, June 2020.

[31] L. Liu, Y. Zhou, V. Garcia, L. Tian and J. Shi, "Load aware joint CoMP clustering and inter-cell resource Scheduling in heterogeneous ultra dense cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2741-2755, March 2018.

[32] M. Peng, Y. Sun, X. Li, Z. Mao and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 2282-2308, Q3 2016.

[33] A. Alimi, A. L. Teixeira and P. P. Monteiro, "Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 1, pp. 708-769, Q1 2018.

[34] M. Jaber, M. A. Imran, R. Tafazolli and A. Tukmanov, "5G backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743-1766, 2016.

[35] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 358-380, Q1 2015.

[36] L. Gavrilovska, V. Rakovic, and D. Denkovski, "From cloud RAN to open RAN". *Wireless Pers Commun.*, Vol. 113, pp. 1523–1539, March 2020.

[37] M. Wani, M. Kretschmer, B. Schröder, A. Grebe and M. Rademacher, "Open RAN: A concise overview," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 13-28, 2025.

[38] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[39] B. M. Zaidel, S. Shamai, and S. Verdú, "Multicell uplink spectral efficiency of coded DS-CDMA with random signatures," *IEEE J. Sel. Areas Commun*., pp. 1556–69, August 2001.

[40] J. G. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Commun. Mag*., pp. 19–29, April 2005.

[41] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," *Proc. IEEE Veh. Tech. Conf*., pp. 1745–9, May 2001.

[42] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE J. Sel. Areas Commun*., pp. 802–11, June 2003.

[43] S. A. Jafar, G. J. Foschini, and A. J. Goldsmith, "PhantomNet: Exploring optimal multicellular multiple antenna systems," *J. Appl. Signal Proc*. (EURASIP), pp. 591–605, May 2004.

[44] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, pp. 439–41, May 1983.

[45] E. H. Choi, W. Choi, and J. Andrews, "Throughput of the lx EV-DO system with various scheduling algorithms," *Proc. Internat. Sympos. Spread Spec. Tech. Appl*., pp. 359–63, August 2004.

[46] D. Wu and R. Negi, "Downlink scheduling in a cellular network for quality-of-service assurance," *IEEE Trans. Veh. Tech*., pp. 1547–57, September 2004.

[47] L. Xu, X. Shen, and J. W. Mark, "Dynamic fair scheduling with QoS constraints in multimedia wideband CDMA cellular networks," *IEEE Trans. Wireless Commun*., pp. 60–73, January 2004.

[48] X. Qiu, L. Chang, Z. Kostic, T. M. Willis, N. Mehta, L. G. Greenstein, K. Chawla, J. F. Whitehead, and J. Chuang, "Some performance results for the downlink shared channel in WCDMA," *Proc. IEEE Internat. Conf. Commun*., pp. 376–80, April 2002.

[49] A. C. Varsou and H. V. Poor, "HOLPRO: A new rate scheduling algorithm for the downlink of CDMA networks," *Proc. IEEE Veh. Tech. Conf.*, pp. 948–54, September 2000.

[50] E. Villier, P. Legg, and S. Barrett, "Packet data transmissions in a W-CDMA network – Examples of uplink scheduling and performance," *Proc. IEEE Veh. Tech. Conf.*, pp. 2449–53, May 2000.

[51] L. Qian and K. Kumaran, "Uplink scheduling in CDMA packet-data systems," *Proc. IEEE Infocom Conf.*, pp. 292–300, March 2003.

[52] L. Qian and K. Kumaran, "Scheduling on uplink of CDMA packet data network with successive interference cancellation," *Proc. IEEE Wireless Commun. Network Conf.*, pp. 1645–50, March 2003.

[53] A. Mamane, M. Fattah, M. E. Ghazi, M. E. Bekkali, Y. Balboul and S. Mazer, "Scheduling Algorithms for 5G Networks and Beyond: Classification and Survey," *IEEE Access*, vol. 10, pp. 51643-51661, 2022.

[54] H. Boche and M. Wiczanowski, "Queueing theoretic optimal scheduling for multiple input multiple output multiple access channel," *Proc. IEEE Internat. Sympos. Signal Proc. Inform. Tech.*, pp. 576–9, December 2003.

[55] K.-N. Lau, "Analytical framework for multiuser uplink MIMO space-time scheduling design with convex utility functions," *IEEE Trans. Wireless Commun.*, pp. 1832–43, September 2004.

[56] L. F. Chang, X. Qiu, K. Chawla, and C. Jian, "Providing differentiated services in EGPRS through radio resource management," *Proc. IEEE Internat. Conf. Commun.*, pp. 2296–2301, June 2001.

[57] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems-A comprehensive survey," *IEEE Pers. Commun. Mag.*, pp. 10–31, June 1996.

[58] D. C. Cox, "Wireless network access for personal communications," *IEEE Commun. Mag.*, pp. 96–115, December 1992.

[59] R. J. McEliece and K. N. Sivarajan, "Performance limits for channelized cellular telephone systems," *IEEE Trans. Inform. Theory*, pp. 21–4, January 1994.

[60] D. Everitt and D. Manfield, "Performance analysis of cellular mobile communication systems with dynamic channel assignment," *IEEE J. Sel. Areas Commun.*, pp. 1172–81, October 1989.

[61] J. Zander and H. Eriksson, "Asymptotic bounds on the performance of a class of dynamic channel assignment algorithms," *IEEE J. Sel. Areas Commun.*, pp. 926–33, August 1993.

[62] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, New York, 1979.

[63] D. Kunz, "Channel assignment for cellular radio using neural networks," *IEEE Trans. Veh. Tech.*, pp. 188–93, February 1991.

[64] R. Mathar and J. Mattfeldt, "Channel assignment in cellular radio networks," *IEEE Trans. Veh. Tech.*, pp. 647–56, November 1993.

[65] A. Lozano and D. C. Cox, "Distributed dynamic channel assignment in TDMA mobile communication systems," *IEEE Trans. Veh. Tech.*, pp. 1397–1406, November 2002.

[66] L. Chen, U. Yoshida, H. Murata, and S. Hirose, "Dynamic timeslot allocation algorithms suitable for asymmetric traffic in multimedia TDMA/TDD cellular radio," *Proc. IEEE Veh. Tech. Conf.*, pp. 1424–8, May 1998.

[67] Y. Hara, T. Nabetani, and S. Hara, "Performance evaluation of cellular SDMA/TDMA systems with variable bit rate multimedia traffic," *Proc. IEEE Veh. Tech. Conf.*, pp. 1731–4, October 2001.

[68] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Tech.*, pp. 57–62, February 1992.

[69] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Tech.*, pp. 641–6, November 1993.

[70] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Trans. Commun.*, pp. 226–8, Febmary–April 1994.

[71] S. A. Grandhi, R. D. Yates, and D. J. Goodman, "Resource allocation for cellular radio systems," *IEEE Trans. Veh. Tech.*, pp. 581–7, August 1997.

[72] A. Lozano and D. C. Cox, "Integrated dynamic channel assignment and power control in TDMA mobile wireless communication systems," *IEEE J. Sel. Areas Commun.*, pp. 2031–40, November 1999.

[73] R. Veronesi, V. Tralli, J. Zander, and M. Zorzi, "Distributed dynamic resource allocation with power shaping for multicell SDMA packet access networks," *Proc. IEEE Wireless Commun. Network Conf.*, pp. 2515–20, March 2004.

[74] F. Berggren, S.-L. Kim, R. Jantti, and J. Zander, "Joint power control and intracell scheduling of DS-CDMA nonreal time data," *IEEE J. Sel. Areas Commun.*, pp. 1860–70, October 2001.

[75] H. C. Akin and K. M. Wasserman, "Resource allocation and scheduling in uplink for multimedia CDMA wireless systems," *IEEE/Sarnoff Sympos. Adv. Wired Wireless Commun.*, pp. 185–8, April 2004.

[76] T. H. Hu and M. M. K. Liu, "A new power control function for multirate DS-CDMA systems," *IEEE Trans. Commun*. pp. 896–904, June 1999.

[77] S.-J. Oh, D. Zhang, and K. M. Wasserman, "Optimal resource allocation in multiservice CDMA networks," *IEEE Trans. Wireless Commun.*, pp. 811–21, July 2003.

[78] D. Ayyagari and A. Ephremides, "Optimal admission control in cellular DS-CDMA systems with multimedia traffic," *IEEE Trans. Wireless Commun.*, pp. 195–202, January 2003.

[79] B. Agarwal, M. A. Togou, M. Marco, and G. -M. Muntean, "A comprehensive survey on radio resource management in 5G HetNets: Current solutions, future trends and open issues," *IEEE Communications Surveys Tutorials*, vol. 24, no. 4, pp. 2495-2534, 2022.

[80] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wang, and T. Q. S. Quek, "Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3481-3493, June 2015.

[81] F. Tang, Y. Zhou, and N. Kato, "Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5G HetNet," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 12, pp. 2773-2782, Dec. 2020.

[82] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 1, pp. 336-361, First Quarter 2013,

[83] S. V. Hanly and P. Whiting, "Information theory and the design of multi-receiver networks," *Proc. Internat. Sympos. Spread Spec. Tech. Appl.*, pp. 103–6, November 1992.

[84] A. Wyner, "Shannon-theoretic approach to a Gaussian cellular," *IEEE Trans. Inform. Theory*, pp. 1713–27, November 1994.

[85] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[86] E. C. van der Meulen, "Some reflections on the interference channel," in R. E. Blahut, D. J. Costello, and T. Mittelholzer, Eds., *Communications and Cryptography: Two Sides of One Tapestry*, pp. 409–21, Kluwer, Boston, 1994.

[87] M. H. M. Costa and A. A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inform. Theory*, pp. 710–11, September 1987.

[88] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels: Part I," *IEEE Trans. Inform. Theory*, pp. 1877–94, November 1997.

[89] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels: Part II," *IEEE Trans. Inform. Theory*, pp. 1895–1911, November 1997.

[90] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Trans. Veh. Tech.*, pp. 1047–66, July 1999.

[91] M. F. Tariz and A. Nix, "Area spectral efficiency of a channel adaptive cellular mobile radio system in a correlated shadowed environment," *Proc. IEEE Veh. Tech. Conf.*, pp. 1075–9, May 1998.

Chapter 16

Ad Hoc Wireless Networks

Ad hoc wireless networks consist of wireless nodes that self-configure to form a network, as shown in Figure 16.1. In contrast to cellular networks, ad hoc wireless networks lack an inherent infrastructure. Thus, all control and networking tasks are performed by the wireless nodes, typically using distributed protocols. Ad hoc wireless networks typically consists of heterogeneous nodes with some node pairs directly connected and other node pairs connected via one or more intermediate wireless relays. Ad hoc wireless networks can be deployed quicker, cheaper, and more easily than networks requiring infrastructure. They are also easily reconfigured if one or more nodes go down, in contrast to a cellular or Wi-Fi network which cannot quickly reconfigure to the loss of a base station or access point. The performance of ad hoc wireless networks is generally inferior to that of infrastructure-based networks such as cellular and Wi-Fi due to the need for multihop routing and the lack of central control and optimization. As discussed in Chaper 1.2, cellular networks may deploy relays and Wi-Fi networks may deploy multiple access points that are connected wirelessly in a mesh Wi-Fi configuration. While these networks entail multihop routing, neither are considered to be an ad hoc wireless network due to their inherent infrastructure.

This chapter first describes the design principles and protocols of ad hoc wireless networks through the concept of protocol layering, including the design tradeoffs and challenges for each layer. The next section summarizes the different routing protocols ad hoc wireless networks employ to route data between network nodes. Fundamental capacity limits and scaling laws for these networks are outlined next, followed by a discussion of the unique design challenges inherent to energy-constrained ad hoc wireless networks. The chapter concludes with an overview of the primary applications for ad hoc wireless networks, including the unique design requirements for each application.

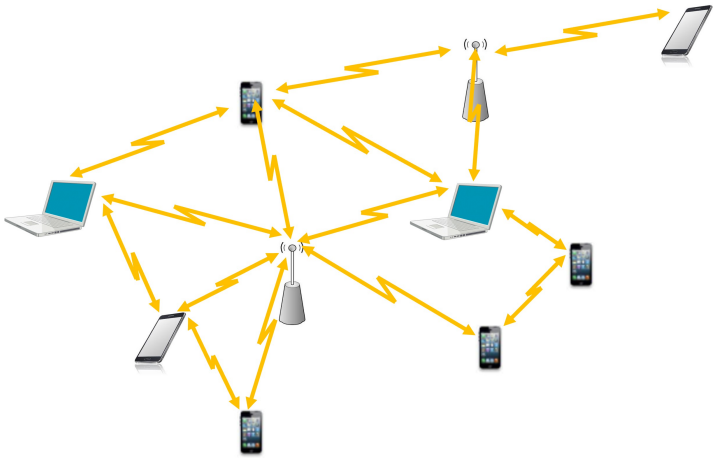


Figure 16.1: Ad hoc wireless network.

16.1 Design Principles

The most fundamental aspect of an ad hoc wireless network is its lack of infrastructure, and most design principles and protocols stem from this characteristic. The lack of infrastructure inherent to ad hoc wireless networks is best illustrated by contrast with the most prevalent wireless networks: cellular systems and wireless LANs. Cellular systems divide the geographic area of interest into cells, and mobiles within a cell communicate with a base station in the cell center that is connected to a backbone wired network. Thus, there is no peer-to-peer communication between mobiles. All communication is via the base station through single-hop routing. The base stations and backbone network perform all networking functions, including authentication, call routing, and handoff. Most wireless LANs have a similar, centralized, single-hop architecture: mobile nodes communicate directly with a centralized access point that is connected to the backbone Internet, and the access point performs all networking and control functions for the mobile nodes.[1] In contrast, an ad hoc wireless network has peer-to-peer communication, networking and control functions that are distributed among all nodes, and routing that can exploit intermediate nodes as relays.

Ad hoc wireless networks can form an infrastructure or node hierarchy, either permanently or dynamically. For example, many ad hoc wireless networks form a backbone infrastructure from a subset of nodes in the network to improve network reliability, scalability, and capacity [26]. If a node in this backbone subset leaves the network, the backbone can be reconfigured. Similarly, some nodes may be chosen to perform as base stations for neighboring nodes [14]. Thus, ad hoc wireless networks may create structure to improve network performance, but such structure is not a fundamental design requirement of the network.

A lack of canonical structure is quite common in wired networks. Indeed, most metropolitan area networks (MANs) and wide area networks (WANs), including the Internet, have an ad hoc structure. However, the broadcast nature of the radio channel introduces characteristics in ad hoc wireless networks that are not present in their wired counterparts. In particular, with sufficient transmit power any node can transmit a signal directly to any other node. For a fixed transmit power, the link signal-to-interference-plus-noise power ratio (SINR) between two communicating nodes will typically decrease as the distance between the nodes increases, and it will also depend on the signal propagation and interference environment. Moreover, this link SINR varies randomly over time owing to fading of the signal and interference. Link SINR determines the communication performance of the link: the data rate and associated probability of packet error or bit error rate (BER) that can be supported on the link. Links with very low SINRs are not typically used due to their extremely poor performance, leading to partial connectivity among all nodes in the network, as shown in Figure 16.1. However, link connectivity is not a binary decision, as nodes can adapt to the SINR using adaptive modulation or change it using power control. The different SINR values for different links are illustrated by the different line widths in Figure 16.1. Thus, in theory, every node in the network can transmit data directly to any other node. However, this may not be feasible if the nodes are separated by a large distance, and direct transmission even over a relatively short link may have poor performance or cause much interference to other links. Network connectivity also changes as nodes enter and leave the network, and this connectivity can be controlled by adapting the transmit power of existing network nodes to the presence of a new node [27].

The flexibility in link connectivity that results from varying link parameters such as power and data rate has major implications for routing. Nodes can send packets directly to their final destination via single-hop routing as long as the link SINR is above some minimal threshold. However, the SINR can be quite poor under single-hop routing, and this routing method also causes excessive interference to surrounding nodes. In most ad hoc wireless networks, packets are forwarded from source to destination through intermediate relay nodes. Since path loss causes an exponential decrease in received power as a function of distance, intermediate relays can greatly reduce the total transmit power (the sum of transmit power at the source and all relays) needed for end-to-end packet

[1] The 802.11 wireless LAN standard does include ad hoc network capabilities, but this component of the standard is rarely used.

transmission. Multihop routing using intermediate relay nodes is a key feature of ad hoc wireless networks: it allows for communication between geographically dispersed nodes and facilitates the scalability and decentralized control of the network. However, it is much more challenging to support high data rates and low delays over a multihop wireless channel than over the single-hop wireless channels inherent to cellular systems and wireless LANs. This is one of the main difficulties in using an ad hoc wireless network to support applications (e.g. video) that require a high data rate and low delay.

Scalability is required for ad hoc wireless networks with a large number of nodes. The key to scalability lies in the use of distributed network control algorithms, which adjust local performance to account for local conditions. By forgoing the use of centralized information and control resources, protocols can scale as the network grows because they rely on local information only. Work on protocol scalability in ad hoc wireless networks has mainly focused on self-organization [28, 29], distributed routing [30], mobility management [26], and security [31]. Note that distributed protocols often consume a fair amount of energy in local processing and message exchange; this is analyzed in detail for security protocols in [32]. Thus, interesting trade-offs arise concerning how much local processing should be done versus transmitting information to a centralized location for processing. This trade-off is particularly apparent in sensor networks, where nodes close together have correlated data and also coordinate in routing that data through the network. Most experimental work on scalability in ad hoc wireless networks has focused on relatively small networks, fewer than a hundred nodes. Many ad hoc network applications, especially sensor networks, could have hundreds to thousands of nodes or even more. The ability of existing wireless network protocols to scale to such large network sizes remains unclear.

Energy constraints are another big challenge in ad hoc wireless network design [33]. These constraints arise in wireless (e.g. sensor) network nodes powered by batteries that cannot be recharged. Hard energy constraints significantly affect network design considerations. First, there is no longer a notion of data rate, since only a finite number of bits can be transmitted at each node before the battery dies. There is also a trade-off between the duration of a bit and energy consumption: sending bits more slowly conserves transmit energy but may increase circuit energy [34]. Standby operation can consume significant energy, so sleep modes must be employed for energy conservation – but then having nodes go to sleep can complicate network control and routing. In fact, energy constraints affect almost all of the network protocols in some manner, so energy consumption must be optimized over all aspects of the network design.

16.1.1 Protocol Layers

Protocol layering is a common abstraction in network design. Layering provides design modularity for network protocols, which facilitates standardization and implementation. Unfortunately, the layering paradigm does not work well in ad hoc wireless networks, where many protocol design issues are intertwined. In this section we describe protocol layering as it applies to ad hoc wireless networks as well as the interactions between protocol layers, which motivate the need for cross-layer design.

An international standard called the open systems interconnect (OSI) model was developed as a framework for protocol layering in data networks. The OSI model divides the required functions of the network into seven layers: the application layer, presentation layer, session layer, transport layer, network layer, data link control layer, and physical layer. Each layer is responsible for a separate set of tasks, with a fixed interface between layers to exchange control information and data. The basic premise behind the OSI model is that the protocols developed at any given layer can interoperate with protocols developed at other layers without regard to the details of the protocol implementation. For example, the application layer need not consider how data is routed through the network, or what modulation and coding technique is used on a given link. The set of protocols associated with all layers is referred to as the *protocol stack* of the network. Details of the OSI model and the functionality associated with each of its layers are given in [35, Chap. 1.3].

The Internet has driven the actual implementation of layering, which is built around the transmission control

protocol (TCP) for the transport layer and the Internet protocol (IP) for routing at the network layer. Thus, in most networks the OSI layering model has been replaced by a five-layer model, also called the TCP/IP model, that is defined by the main functionality of the TCP and IP protocols. The five layers consist of the application layer, transport layer, network layer, access layer, and physical layer. These layers are illustrated in Figure 16.2, along with their primary functions in ad hoc wireless networks. These functions will be described in more detail below. Note that power control sits at two layers, the physical and access layers, and is part of resource allocation at the network layer as well [36].

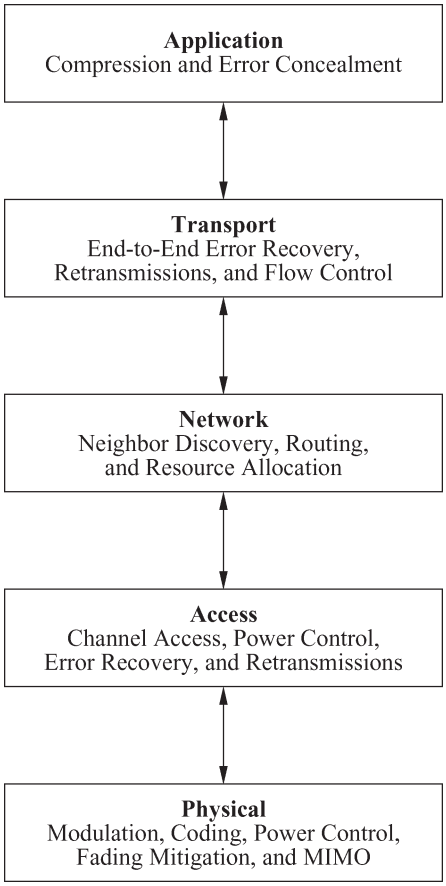


Figure 16.2: Five-layer model for network protocol design.

Thus, power control spans multiple layers of the protocol stack, as discussed in more detail below. Most ad hoc wireless network designs do not use the IP protocol for routing, since routing through a wireless network is much different than for the Internet. Moreover, the addressing and subnetting in the IP protocol is not well suited to ad hoc wireless networks. Transport protocols do not necessarily use TCP either. However, the five-layer model is a common abstraction for the modular design of protocol layers in wireless networks.

The layering principle for protocol design works reasonably well in wired networks like the Internet, where the data rates associated with the physical layer can exceed gigabits per second and packets are rarely lost. However, even in this setting, layering makes it difficult to support high–data-rate applications with hard delay constraints, such as video or even voice. Wireless networks can have very low physical layer data rates and very high packet and bit error probabilities. In this setting, protocol layering can give rise to tremendous inefficiencies and also precludes exploiting interactions between protocol layers for better performance. Crosslayer design considers multiple layers of the protocol stack together, either in terms of a joint design or in information exchange between

the layers. Cross-layer design can exhibit tremendous performance advantages over a strictly layered approach. We now describe the layers of the five-layer model and their functionality in wireless networks. We then discuss the basic principles of cross-layer design and the performance advantages of this approach over strict layering.

16.1.2 Physical Layer Design: Point to Point Communications

The physical layer deals primarily with transmitting bits over a point-to-point wireless link, so it is also referred to as the *link layer*. Chapters 5–13 comprehensively covered the design trade-offs associated with the physical layer, including modulation, coding, diversity, adaptive techniques, MIMO, equalization, multicarrier modulation, and spread spectrum. However, the design trade-offs for a link that is part of an ad hoc wireless network have an impact on protocol layers above the physical layer. In fact, few aspects of the physical layer design in an ad hoc wireless network do not affect in some way the protocols associated with the higher layers. We now give some examples of this interaction for physical layer design choices related to packet error rate, multiple antennas, and power control.

In most wireless ad hoc networks, bits are packetized for transmission as described in Section 14.3. The design choices at the physical layer along with the channel and interference conditions determine the link packet error rate (PER). Many access layer protocols retransmit packets received in error, so PER based on the physical layer design affects the retransmission requirements at the access layer. Similarly, as described in Section 14.10, multiple antennas give rise to a multiplexing-diversity-directionality trade-off: the antennas can be used to increase the data rate on the link, to provide diversity in fading so that average BER is reduced, or to provide directionality to reduce fading and the interference a signal causes to other signals. Diversity gain will reduce PER, leading to fewer retransmissions. Multiplexing will increase the link rate, which reduces congestion and delay on the link and benefits all multihop routes using that link. Directionality reduces interference to other links, thereby improving their performance. Thus, the best use of multiple antennas in an ad hoc wireless network clearly transcends just the physical layer; in fact, it simultaneously affects the physical, access, network, and transport layers.

The transmit power of a node at the physical layer also has a broad impact across many layers of the protocol stack. Increasing transmit power at the physical layer reduces PER, thereby affecting the retransmissions required at the access layer. In fact any two nodes in the network can communicate directly with sufficiently high transmit power, so this power drives link connectivity. However, a high transmit power from one node in the network can cause significant interference to other nodes, thereby degrading their performance and possibly breaking their connectivity on some links. In particular, link performance in an ad hoc wireless network is driven by SINR, so the transmit power of all nodes has an impact on the performance of all links in the network. Broadly speaking, the transmit power coupled with adaptive modulation and coding for a given node defines its "local neighborhood" – the collection of nodes that it can reach in a single hop – and thus defines the context in which access, routing, and other higher-layer protocols operate. Therefore, the transmit power of all nodes in the network must be optimized with respect to all layers that it affects. As such, it is a prime motivator for a cross-layer design.

16.1.3 Access Layer Design: Multiple Access

The access layer controls how different users share the available spectrum and ensures successful reception of packets transmitted over this shared spectrum. Allocation of signaling dimensions to different users is accomplished through either multiple access or random access, and a detailed discussion of these access techniques can be found in Sections 14.2 and 14.3. Multiple access divides the signaling dimensions into dedicated channels via orthogonal or nonorthogonal channelization methods. The most common of these methods are TDMA, FDMA, and CDMA. The access layer must also provide control functionality in order to assign channels to users and to deny access to users when they cannot be accommodated by the system. In random access, channels are assigned to active users dynamically, and in multihop networks these protocols must contend with hidden and exposed terminals. The most common random access methods are different forms of ALOHA, carrier-sense multiple access

(CSMA), and scheduling (see Section 14.3). These random access methods incorporate channel assignment and access denial into their protocols.

As discussed in the prior section, transmit power associated with a single node affects all other nodes. Thus, power control across all nodes in the network is part of the access layer functionality. The main role of power control is to ensure that SINR targets can be met on all links in the network. This is often unfeasible, as discussed in more detail below. The power control algorithms described in Sections 14.5 and 15.3.2 for meeting SINR targets in multiple access and cellular systems, respectively, can be extended to ad hoc networks as follows. Consider an ad hoc wireless network with K nodes and N links between different transmitter–receiver pairs of these nodes.[2] The SINR on link k is given by

$$\gamma_k = \frac{g_{kk}P_k}{n_k + \rho\sum_{j\neq k} g_{kj}P_j}, \quad k,\, j \in \{1, 2, \ldots, N\}, \tag{16.1}$$

where $g_{kj} > 0$ is the channel power gain from the transmitter of the jth link to the receiver of the kth link, P_k is the power of the transmitter on the kth link, n_k is the noise power of the receiver on the kth link, and ρ is the interference reduction due to signal processing (i.e., $\rho \approx 1/G$ for CDMA with processing gain G and $\rho = 1$ in TDMA). Suppose that the kth link requires an SINR of γ_k^* as determined, for example, by the connectivity and data-rate requirements for that link. Then the SINR constraints for all links can be represented in matrix form as $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ with $\mathbf{P} > 0$, where $\mathbf{P} = (P_1, P_2, \ldots, P_N)^T$ is the vector of transmit powers associated with the transmitters on the N links,

$$\mathbf{u} = \left(\frac{\gamma_1^* n_1}{g_{11}}, \frac{\gamma_2^* n_2}{g_{22}}, \ldots, \frac{\gamma_N^* n_N}{g_{NN}}\right)^T \tag{16.2}$$

is the vector of noise powers scaled by the SINR constraints and channel gains, and $\mathbf{F}$ is a matrix with

$$F_{kj} = \begin{cases} 0 & k = j, \\ \gamma_k^* g_{kj}\rho/g_{kk} & k \neq j, \end{cases} \tag{16.3}$$

where $k, j \in \{1, 2, \ldots, N\}$. As in the uplink and cellular power control problems described in Section 14.5 and Section 15.3.2, if the Perron–Frobenius (maximum modulus) eigenvalue of $\mathbf{F}$ is less than unity then there exists a vector $\mathbf{P} > 0$ (i.e., $P_k > 0$ for all k) such that the SINR requirements of all links are satisfied, with $\mathbf{P}^* = (\mathbf{I} - \mathbf{F})^{-1}\mathbf{u}$ the Pareto optimal solution [37, 38, 39]. Thus $\mathbf{P}^*$ meets the SINR requirements with the minimum transmit power on all links. Moreover, a distributed iterative power control algorithm, where the transmitter on the kth link updates its transmit power at time $i + 1$ as

$$P_k(i+1) = \frac{\gamma_k^*}{\gamma_k(i)} P_k(i), \tag{16.4}$$

can be shown [38, 39] to converge to the optimal solution $\mathbf{P}^*$. This is a simple distributed algorithm for power control in an ad hoc wireless network, since it requires only that the SINR of the kth link be known to the transmitter on that link. Then, if this SINR is below its target the transmitter increases its power, and if it is above this target the transmitter decreases the power. It is quite remarkable that such a simple distributed algorithm converges to a globally optimal power control. However, when the channel gains are not static, SINR constraints can no longer be met with certainty, and it is much more difficult to develop distributed power control algorithms that meet a desired performance target [40]. In particular, the algorithm described by (16.4) can exhibit large fluctuations in link SINR when the channel gains vary over time. More importantly, it is often impossible to meet the SINR constraints of all

[2] As noted previously, all nodes can communicate with all other nodes, so there are $K(K-1)$ links for a network with K nodes. However, we assume that only $N \leq K(K-1)$ of these are used, so we need only consider the SINR on these N links.

nodes simultaneously even when the link gains are static, owing to the large number of interferers and the range of channel gains associated with all signals in the network. When the SINR constraints cannot be met, the distributed power control algorithm will diverge such that all nodes transmit at their maximum power and still cannot meet their SINR constraints. This is obviously an undesirable operational state, especially for energy-constrained nodes.

In [41], the distributed power control algorithm assuming static link gains is extended to include distributed admission control. The admission control provides protection for existing link SINR targets when a new user enters the system. In this scheme the active links have a slightly higher SINR target than needed. This buffer is used so that, when a new user who attempts to access the system transmits at a low power level, he will not cause the active links to fall below their minimum SINR targets. The new user gradually ramps up his power and checks if he gets closer to his SINR target as a result. If the new user can be accommodated by the system without violating the SINR constraints of existing links, then the distributed algorithm with this gradual ramp-up will eventually converge to a new $\mathbf{P}^*$ that satisfies the SINR constraints of the new and existing links. However, if the new user cannot be accommodated, his gradual ramp-up will not approach his required SINR and he will eventually be forced to leave the system. Note that the criteria for denial of access is difficult to optimize in time-varying channels with distributed control [40]. These ideas are combined with transmission scheduling in [42, 43] to improve power efficiency and reduce interference. The power control algorithm can also be modified to take into account delay constraints [27, 44]. However, delay constraints are associated with the full multihop route of a packet, so power control should be coordinated with network layer protocols to ensure delay constraints are met on the end-to-end route.

The access layer is also responsible for retransmissions of packets received in error over the wireless link, often referred to as the automatic repeat request (ARQ) protocol. Specifically, data packets typically have an error detection code that is used by the receiver to determine if one or more bits in the packet were corrupted and cannot be corrected. For such packets, the receiver will usually discard the corrupted packet and inform the transmitter via a feedback channel that the packet must be retransmitted. However, rather than discarding the packet, the access layer can save it and use a form of diversity to combine the corrupted packet with the retransmitted packet for a higher probability of correct packet reception. Alternatively, rather than retransmitting the original packet in its entirety, the transmitter can just send some additional coded bits to provide a stronger error correction capability for the receiver to correct for the corrupted bits in the original packet. This technique is called *incremental redundancy*, since the transmitter need only send enough redundant bits to correct for the bits corrupted in the original packet transmission. Diversity and incremental redundancy methods have been shown to substantially improve throughput in comparison with simple retransmissions [45].

16.1.4 Network Layer Design: Discovery, Routing, and Resource Allocation

The network layer is responsible for establishing and maintaining end-to-end connections in the network. For every network node to be able to send data to every other node, the network must be *fully connected*, whereby every node in the network can communicate with every other node, either directly or by multihop routing through intermediate nodes.[3] The main functions of the network layer in an ad hoc wireless network are neighbor discovery, routing, and dynamic resource allocation. Neighbor discovery is the process by which a node discovers its neighbors when it first enters the network. Routing is another key function of the network layer: the process of determining how packets are routed through the network from their source to their final destination or destinations. Routing through intermediate nodes is typically done by relaying, although network coding and similar techniques that exploit multiuser diversity can also be used. Dynamic resource allocation dictates how network resources such as power and bandwidth are allocated throughout the network, although resource allocation in general occurs at multiple layers of the protocol stack and thus requires cross-layer design.

[3]This definition differs from that in graph theory, where every node in a fully connected graph has an edge to every other node.

NEIGHBOR DISCOVERY AND TOPOLOGY CONTROL

Neighbor discovery is one of the first steps in the initialization of a network with randomly distributed nodes. From the perspective of the individual node, this is the process of determining the number and identity of network nodes with which direct communication can be established, given some maximum power level and minimum link performance requirements (typically in terms of data rate and associated BER). Clearly, the higher the allowed transmit power, the greater the number of nodes in a given neighborhood. Note that different addressing schemes at different protocol layers can complicate neighbor discovery, which must resolve these differences.

Neighbor discovery typically begins with a probe of neighboring nodes using some initial transmit power. If this power is not sufficient to establish a connection with $N \geq 1$ neighbors then transmit power is increased and probing repeated. The process continues until N connections are established or the maximum power $P_{\max}$ is reached. The parameter N is set based on network requirements for minimal connectivity, while $P_{\max}$ is based on the power limitations of each node and the network design. If N and/or $P_{\max}$ is small then the network may form in a disconnected manner, with small clusters of nodes that communicate together but cannot reach other clusters. This is illustrated in Figure 16.3, where the dashed circles centered around a node indicate the neighborhood within which it can establish connections with other nodes. If N and $P_{\max}$ are large then – although the network is typically fully connected – many nodes are transmitting at higher power than necessary for full network connectivity, which can waste power and increase interference. Once the network is fully connected, a more sophisticated distributed power control algorithm such as (16.4) can be activated to meet target SINR levels on all links with minimal transmit power. Alternatively, power control can be used to create a desired topology [8].

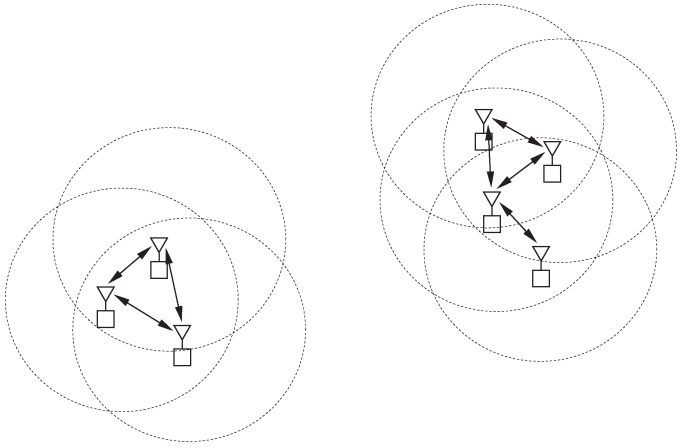


Figure 16.3: Disconnected network.

The exact number of neighbors that each node requires in order to establish a fully connected network depends on the exact network configuration and channel characteristics, but it is generally on the order of six to eight for randomly distributed immobile nodes with channels characterized by path loss alone [2, 35]. The number required for full connectivity under more general assumptions is analyzed in [46, 47, 48]. As node mobility increases, links typically experience large gain variations due to fading. These variations can make it harder for the network to stay fully connected at all times unless the nodes can increase their transmit powers to compensate for instantaneous fading. If the data is tolerant of delay then fading can actually improve network connectivity because it provides network diversity [49]. As network density decreases, network connectivity typically suffers [47, 50, 51, 52]. Connectivity is also heavily influenced by the ability to adapt at the physical layer such parameters as rate, power, and coding, since communication is possible even on links with low SINR if these parameters are adapted.

ROUTING

Once the network topology is established through neighbor discovery and topology control, each node in the network can send data to other nodes to which it is connected. These connections are either single hop (direct transmission) or multihop (using one or more intermediate relays). There are often multiple possible routes to send data between network nodes. The routing protocol in an ad hoc wireless (or wired) network establishes how data is routed from a source node to its final destination node or nodes along one or more of these possible routes. These routes are ideally selected to maximize overall network performance. However, given the complexity of defining and optimizing overall network performance, routes are often selected to minimize a cost function related to performance, such as the number of hops, end-to-end delay, energy consumption, or a weighted combination of different parameters. This minimum cost can be computed centrally for all routes in the network, or in a distributed manner for individual source-destination pairs. There is broad and extensive work spanning many decades on routing protocols for ad hoc wireless networks. In this section we describe the main considerations and paradigms for these routing protocols. Common wireless network routing protocols are described in Section 16.5.

Routing in wireless ad hoc networks is significantly more challenging than in wired networks due to multiple factors. Wired networks have relatively static topologies and channel conditions. Link states in such networks change primarily due to congestion, and thus wired routing protocols work to minimize congestion on any given route or link. By contrast, in wireless networks routes need to be dynamically reconfigured when the network topology changes due to nodes entering and leaving the network or changes in channel conditions between nodes. In addition, for wireless networks a radio broadcast causes interference to nearby radio receivers associated with different routes, whereas in wired networks transmissions do not interfere with each other. In addition, a radio transmission can simultaneously send a data stream to all receivers within range, whereas in wired networks transmitting data to multiple receivers requires special routing. Despite these differences, the majority of routing protocols for ad hoc wireless networks use similar techniques as those for wired networks.

There are three main paradigms in ad hoc wireless network routing - unicast, broadcast, and multicast, as illustrated in Figure 16.4. In unicast routing a data stream is sent from one source to one destination. Unicast routing is used the data from the source is desired by only one destination. In some communication paradigms, data from a source is desired by all nodes in the network, for example in video broadcasting or to communication network control information. For this communication paradigm, broadcast routing is used, whereby a data stream is sent from one source to all nodes in the network. Broadcast routing is a special case of multicast routing, whereby a data stream is sent from one source to a subset of all nodes in the network. Multicast routing becomes broadcast routing when the subset of nodes equals the entire set of nodes.

It is typically the case for all three routing paradigms that a data stream traverses a single path to its final destination or destinations. However, some networks utilize multipath routing, where a packet is duplicated on a few end-to-end paths between its source and destination. Since it is unlikely that the duplicate packets are lost or significantly delayed on all paths simultaneously, the packet has a high probability of reaching its final destination with minimal delay on at least one of the paths [53]. On the other hand, it is more complex to compute and implement multipath routing and the duplication of packets increases overall network congestion while reducing available bandwidth. In addition, duplicate packets and packets arriving out of order complicate recreation of the full data stream at its destination.

RESOURCE ALLOCATION AND FLOW CONTROL

A routing protocol dictates the route a packet should follow from a source node to its destination. When the routing optimization is based on minimum congestion or delay, routing becomes intertwined with flow control, which typically sits at the transport layer. If the routing algorithm sends too much data over a given link, then this link becomes congested and so the routing algorithm should find a different route to avoid this link. Moreover, the

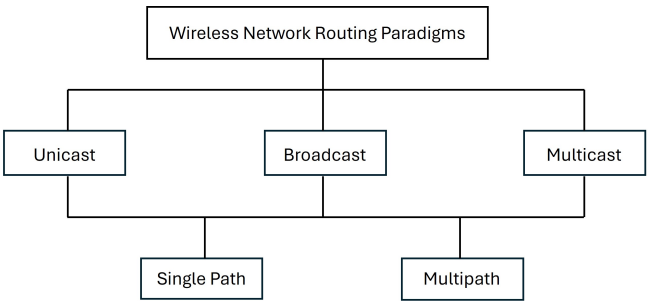


Figure 16.4: Wireless Network Routing Paradigms

delay associated with a given link is a function of the link data rate or capacity: the higher the capacity, the more data that can flow over the link with minimal delay. Because link capacity depends on the resources allocated to the link – in particular, transmit power and bandwidth – we see that routing, resource allocation, and flow control are all interdependent.

The classic metric for delay on a link from node i to node j, neglecting processing and propagation delay, is

$$D_{ij} = \frac{f_{ij}}{C_{ij} - f_{ij}} \tag{16.5}$$

[35, Chap. 5.4], where f_{ij} is the traffic flow assigned to the link and C_{ij} is its capacity (the maximum flow that the link can support). This formula has its roots in queueing theory and provides a good metric in practice, since the closer the flow is to the maximum data rate on a given link, the more likely the link will become congested and incur delay. Another metric on the link between nodes i and j is the link utilization, given by

$$D_{ij} = \frac{f_{ij}}{C_{ij}}. \tag{16.6}$$

As discussed in [35, Chap. 5.4], this metric has properties that are comparable to those of the delay metric (16.5) and is also a quasi-convex function[4] of both flow and capacity, which allows efficient convex optimization methods to be applied in the routing computations. If the data flows across links in the network are fixed, then the routing algorithm can compute the per-hop cost for each link based on the delay metric (16.5) or the utilization metric (16.6) and use these per-hop costs to find the minimum cost route through the network. The difference between these two metrics is that delay grows asymptotically large as the flow approaches link capacity, whereas link utilization approaches unity. Thus, the delay metric (16.5) has a much higher cost than the utilization metric (16.6) when flow is assigned to a link that is operating at close to its capacity. Once the new route is established, it will change the link flows along that route. In most cases this change in flow will not be large, since the contribution of any one node to overall traffic is not large, but in small to moderately sized networks a demanding application such as video can cause significant self-congestion.

The link metrics (16.5) and (16.6) assume a fixed link capacity C_{ij}. However, this capacity is a function of the SINR on the link and also of the bandwidth allocated to that link. By dynamically allocating network resources such as power and bandwidth to congested links, their capacities can be increased and their delay reduced. However, this may take away resources from other links, thereby decreasing their capacity. These changes in link capacity will in turn change the link metrics used to compute optimal routes and thus ultimately affect the overall performance of the network. Hence, the performance of the network depends simultaneously on routing, flow control, and resource allocation.

[4] A function is *quasi-convex* if the set over which its value is below any given threshold is convex.

The joint optimization of flow control, routing, and resource allocation can be formulated as a convex optimization problem over the flow and communications variables, assuming the cost and capacity functions are convex (or quasi-convex). Interior-point convex optimization methods can then be applied to solve for the optimal design. This approach has been investigated in [81, 82, 83] for both TDMA and CDMA wireless networks in order to minimize power, maximize link utilization, or maximize flow utility through joint routing and resource allocation. Similar ideas using iterative optimization were explored in [84]. The maximum throughput in this setting can lead to highly unfair allocation of system resources [85], although the framework can be modified to include fairness constraints [86].

16.1.5 Transport Layer Design: End-to-End Communication

The transport layer provides the end-to-end functions of error recovery, retransmission, reordering, and flow control. Although individual links provide error detection and retransmissions, these mechanisms are not foolproof. The transport layer provides an extra measure of protection by (i) monitoring for corrupt or lost packets on the end-to-end route and (ii) requesting a retransmission from the original source node if a packet is determined to be lost or corrupted. In addition, packets may arrive out of order owing to multipath routing, delays and congestion, or packet loss and retransmission. The transport layer serves to order packets transmitted over an end-to-end route before passing them to the application layer.

The transport layer also provides flow control for the network, allocating flows associated with the application layer to different routes. The transmission control protocol for the transport layer does not work well in wireless networks, since it assumes all lost packets are due to congestion and invokes congestion control as a result. In wired networks, congestion is the primary reason for packet loss and so the TCP works well; this is why it is used for the transport layer of the Internet. However, in wireless networks packets are mostly lost as a result of channel impairments and node mobility. Invoking congestion control in this case can lead to extreme inefficiency [57, Chap. 11.5]. There has been some limited progress on developing mechanisms to improve TCP performance for wireless networks by providing transport-layer feedback about link failures.

In general, flow control in wireless networks is intricately linked to resource allocation and routing, as described in Section 16.1.4. This interdependency is much tighter in wireless networks than in their wired counterparts. In particular, wired networks have links with fixed capacity, whereas the capacity of a wireless link depends on the interference between links. Traffic flows assigned to a given link will cause interference to other links, thereby affecting their capacity and delay. This interdependency makes it difficult to separate out the functions of flow control, resource allocation, and routing into separate network and transport layers, motivating a cross-layer design encompassing both layers.

16.1.6 Application Layer Design

The application layer generates the data to be sent over the network and processes the corresponding data received over the network. As such, this layer provides compression of the application data along with error correction and concealment. The compression must be loss-less for data applications but can be lossy for video, voice, or image applications, where some loss can be tolerated in the reconstruction of the original data. The higher the level of compression, the less the data rate burden imposed on the network. However, highly compressed data is more sensitive to errors because most of the redundancy has been removed. Data applications cannot tolerate any loss, so packets that are corrupted or lost in the end-to-end transmission must be retransmitted, which can entail significant delay. Voice, video, and image applications can tolerate some errors, and techniques like error concealment or adaptive playback can mitigate the impact of these errors on the perceived quality at the receiving end [87, 88]. Thus, a trade-off at the application layer is data rate versus robustness: higher rates burden the network but enable data transmissions that are more robust with respect to network performance.

The application layer can also provide a form of diversity through multiple description coding (MDC) [89, 90], which is a form of compression whereby multiple descriptions of the data are generated. The original data can be reconstructed from any of these descriptions with some loss, and the more descriptions that are available, the better the reconstruction. If multiple descriptions of the source data are sent through the network then some of these descriptions can be lost, delayed, or corrupted without significantly degrading overall performance. Thus, MDC provides a form of diversity at the application layer to unreliable network performance. Moreover, MDC can be combined with multipath routing to provide cross-layer diversity in both the application description and the routes over which these descriptions are sent [91]. The trade-off is that, for a given data rate, an MDC entails more distortion than a compression technique that is not geared to providing multiple descriptions. This can be viewed as a performance–diversity trade-off: the application sacrifices some level of performance in order to provide robustness to uncertainty in the network.

Many applications require a guaranteed end-to-end data rate and delay for good performance, collectively referred to as QoS. The Internet today – even with high-speed, high-quality fixed communication links – is unable to deliver guaranteed QoS to applications in terms of guaranteed end-to-end rates or delays. For ad hoc wireless networks – with their low-capacity, error-prone time-varying links, mobile users, and dynamic topologies – the notion of being able to guarantee these forms of QoS is simply unrealistic. Therefore, ad hoc wireless network applications must adapt to the time-varying QoS parameters offered by the network. Although adaptivity in the physical, access, and network layers (as described in previous sections) will provide the best possible QoS to the application, this QoS will vary with time as channel conditions, network topology, and user demands change. Applications should therefore adapt to the QoS that is offered. There can also be some negotiation in which users with a higher priority can obtain a better QoS by lowering the QoS of less-important users.

As a simple example, the network may offer the application a rate–delay trade-off curve that is derived from the capabilities of the lower-layer protocols [92]. The application layer must then decide at which point on this curve to operate. Some applications may be able to tolerate a higher delay but not a lower overall rate. Examples include data applications in which the overall data rate must be high yet some latency is tolerable. Other applications (e.g., a distributed control application) might be extremely sensitive to delay but might be able to tolerate a lower rate (e.g., via a coarser quantization of sensor data). Lossy applications like voice or video might exchange some robustness to errors for a higher data rate. Energy constraints introduce another set of trade-offs related to network performance versus longevity. Hence trade-off curves in network design will typically be multidimensional, incorporating rate, delay, robustness, and longevity. These trade-offs will also change with time as the number of users in the network and the network environment change.

16.1.7 Cross-Layer Design

The decentralized control, lack backbone infrastructure, and unique characteristics of wireless links make it difficult to support demanding applications over ad hoc wireless networks, especially applications with high–data-rate requirements and hard delay constraints. The layering approach to wireless (and wired) network design – where each layer of the protocol stack is oblivious to the design and operation of other layers – has not worked well in general, especially under stringent performance requirements. Layering precludes the benefits of joint optimization discussed in prior sections. Moreover, good protocol designs for isolated layers often interact in negative ways across layers, which can significantly degrade end-to-end performance and also make the network extremely fragile to network dynamics and interference. Thus, stringent performance requirements for wireless ad hoc networks can only be met through a cross-layer design. Such a design requires that the interdependencies between layers be characterized, exploited, and jointly optimized. Cross-layer design clearly requires information exchange between layers, adaptivity to this information at each layer, and diversity built into each layer to ensure robustness.

Although cross-layer design can be applied to both wireless and wired networks, wireless ad hoc networks pose unique challenges and opportunities for this design framework owing to the characteristics of wireless chan-

nels. The existence of a link between nodes, which can be used to communicate between the nodes or cause them to interfere with each other, can be controlled by adaptive protocols such as adaptive modulation and coding, adaptive signal processing in space, time, or frequency, and adaptive power control. Since higher-layer protocols (access and routing) depend on underlying node connectivity and interference, adaptivity at the physical layer can be exploited by higher-layer protocols to achieve better performance. At the same time, some links exhibit extreme congestion or fading. Higher-layer protocols can bypass such links through adaptive routing, thereby minimizing delays and bottlenecks that arise as a result of weak links. At the highest layer, information about the throughput and delay of end-to-end routes can be used to change the compression rate of the application or send data over multiple routes via MDCs. Thus, higher-layer protocols can adapt to the status of lower layers.

Adaptation at each layer of the protocol stack should compensate for variations at that layer based on the time scale of these variations. Specifically, variations in link SINR are very fast: on the order of microseconds for fast fading. Network topology changes more slowly, on the order of seconds, while variations of user traffic based on their applications may change over tens to hundreds of seconds. The different time scales of the network variations suggest that each layer should attempt to compensate for variation at that layer first. If adapting locally is unsuccessful then information should be exchanged with higher layers for a broader response to the problem. For example, suppose the SINR on a given link in an end-to-end route is low. By the time this connectivity information is relayed to a higher level of the protocol stack (i.e., the network layer for rerouting or the application layer for reduced-rate compression), the link SINR will most likely have changed. Therefore, it makes sense for each protocol layer to adapt to variations that are local to that layer. If this local adaptation is insufficient to compensate for the local performance degradation then the performance metrics at the next layer of the protocol stack will consequently degrade. Adaptation at this next layer may then correct or at least mitigate the problem that could not be fixed through local adaptation. For example, consider again a low SINR link. Link SINR can be measured quite accurately and quickly at the physical layer. The physical layer protocol can therefore respond to the low SINR by increasing transmit power or the level of error correction coding. This will correct for variations in connectivity due to, for example, multipath flat fading. However, if the weak link is caused by something difficult to correct for at the physical layer – for example, the mobile unit is inside a tunnel – then it is better for a higher layer of the network protocol stack to respond by (say) delaying packet transmissions until the mobile leaves the tunnel. Similarly, if nodes in the network are highly mobile then link characteristics and network topology will change rapidly. Informing the network layer of highly mobile nodes might change the routing strategy from unicast to broadcast in the general direction of the intended user. Ultimately, if the network cannot deliver the QoS requested by the application, then the application must adapt to whatever QoS is available or suffer the degradation associated with requiring network performance that cannot be delivered. It is this integrated approach to adaptive networking – how each layer of the protocol stack should respond to local variations given adaptation at higher layers – that constitutes an adaptive cross-layer protocol design.

Diversity is another mechanism that can be exploited in cross-layer design. Diversity is commonly used to provide robustness to fading at the physical layer. However, the basic premise of diversity can be extended across all layers in the network protocol stack. Cooperative diversity provides diversity at the access layer by using multiple spatially distributed nodes to aid in forwarding a given packet. This provides robustness to packet corruption on any one link. Network layer diversity is inherent to multipath routing, where multiple routes through the network are used to send a single packet. This induces a diversity–throughput trade-off at the network layer that is similar to that described for MIMO systems at the physical layer (see Section 10.5). Specifically, a packet transmitted over multiple routes through the network is unlikely to be dropped or significantly delayed simultaneously on all routes. Thus, the packet dropping probability and average delay is decreased by network diversity. However, such a packet utilizes network resources that could be used to send other packets, thereby reducing overall network throughput. Application layer diversity follows from using MDCs to describe the application data: as long as one of the descriptions is received, the source data can be reproduced (albeit with higher distortion than if the reproduction is

based on all descriptions). Diversity across all layers of the protocol stack, especially when coupled with adaptive cross-layer design, can ensure reliability and good performance over wireless ad hoc networks despite their inherent challenges.

Cross-layer design across multiple protocol layers below the application layer were discussed in the preceding sections. Cross-layer design that includes the application layer along with lower layers is a difficult challenge requiring interdisciplinary expertise, and there is little work addressing this challenge to date. Moreover, there are potential pitfalls to cross-layer design, including increased complexity and reduced architectural flexibility [93]. Even so, the potential performance gains are significant, as illustrated by the cross-layer designs in [94, 95, 96] for video and image transmission in ad hoc wireless networks.

Cross-layer design is particularly important in energy-constrained networks, where each node has a finite amount of energy that must be optimized across all layers of the protocol stack. Energy constraints pose unique challenges and opportunities for cross-layering. Some of these design issues will be discussed in Section 16.4.4.

16.2 Routing Protocols

In this section we describe the main protocols for routing in ad hoc wireless networks, along with their attributes and tradeoffs. A more detailed overview of this broad topic can be found in [54, 55, 56][57, Chap. 5]. The are seven main types of routing protocols - flooding, proactive, reactive, hybrid, hierarchical, location-based, and power-aware - as shown in Figure 16.5. The simplest of these protocols is flooding, where a packet is broadcast to all nodes within receiving range. These nodes also broadcast the packet, and the forwarding continues until the packet reaches its ultimate destination. Flooding has the advantage that it is highly robust to changing network topologies and requires little routing overhead. In fact, in highly mobile networks flooding may be the only feasible routing strategy. The obvious disadvantage is that multiple copies of the same packet traverse through the network, wasting bandwidth and battery power of the transmitting nodes. This disadvantage makes flooding impractical for all but the smallest of networks.

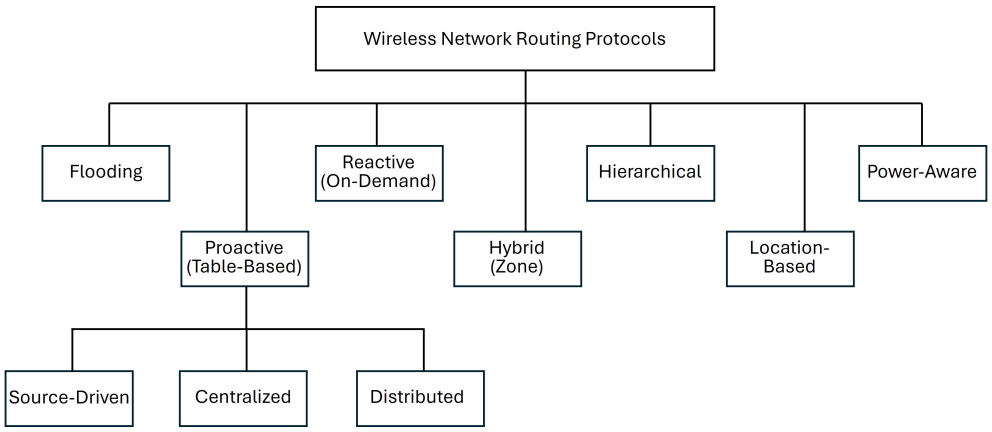


Figure 16.5: Types of Wireless Network Routing Protocols

The opposite philosophy to flooding is proactive route computation. In this approach, information about channel conditions and network topology is used to compute the routing tables for all source-destination node pairs in the network. This method is proactive since routes between all node pairs are computed even if some might never be used. The criterion used to compute each source-destination route depends on the optimization objectives. Common objectives for route optimization include minimum average delay, minimum number of hops, and minimum network congestion [58]. In general these objectives are mapped to the cost associated with each hop along a route.

The minimum cost route between the source and destination is then obtained using classic optimization techniques such as the Bellman–Ford or Dijkstra's algorithm (see [59]). Route optimization to minimize hop count is called *distance vector routing*. The most common examples of distance vector routing in ad hoc wireless networks are the Wireless Routing Protocol (WRP) and Destination-Sequenced Distance-Vector Routing (DSDV). These protocols both aim to minimize hop count on each route, trading off complexity and memory storage for convergence speed. Minimizing the hop count does not necessarily lead to good end-to-end performance when one or more hops in a route have poor link quality. Better end-to-end performance can be obtained by optimizing routes with respect to a more general cost function associated with total hop count as well as the link quality on each hop. This method is called *link state routing*. Link state routing protocols can differ in the link state metrics used, their end-to-end optimization. and their implementation. The most common example of link state routing in wireless ad hoc networks is the Optimized Link State Routing (OLSR) protocol. A survey describing and comparing the most common distance vector and link state routing protocols, including the above examples, can be found in [56].

In centralized route computation, all information about channel conditions and network topology is used by a central controller to compute the routing tables. Although centralized route computation provides the most efficient routing in terms of the optimality condition, under high mobility routing tables quickly become outdated, requiring an enormous amount of overhead to periodically collect local node information and then disseminate the routing information. Centralized route computation, like flooding, is typically used only in very small networks. An alternative to this centralized computation is source-driven routing, where each node obtains connectivity information about the entire network that is then used to calculate the best route from the node to its desired destination. Source-driven routing must also periodically collect network connectivity information, which entails significant overhead.

Distributed route computation is the most common proactive routing procedure in ad hoc wireless networks. In this protocol, nodes send their connectivity information to neighboring nodes and then routes are computed from this local information. In particular, nodes determine the next hop in the route of a packet based on this local information. There are several advantages to distributed route computation. First, the overhead of exchanging routing information with local nodes is minimal. In addition, this strategy adapts quickly to link and connectivity changes. The disadvantages of this strategy are that global routes based on local information are typically suboptimal, and *routing loops* – routes where a packet cycles between the same intermediate nodes without ever reaching its final destination – are often common in the distributed route computation. These loops can be avoided using various strategies. For example, DSDV incorporates sequence numbers as part of the routing tables to avoid such loops [60]. Because of network dynamics, the routing tables obtained by proactive route optimization must be updated at regular intervals, which entails significant overhead, particularly in networks with rapidly changing topologies.

An alternate approach to proactive routing that entails much less overhead is reactive (on-demand) routing. Under this protocol, routes are created only at the initiation of a source node that has traffic to send to a given destination. This eliminates the overhead of maintaining routing tables for routes not currently in use. In this strategy, a source node initiates a route discovery process when it has data to send. This process will determine if one or more routes are available to the destination. The route or routes are maintained until the source has no more data for that particular destination. The advantage of reactive routing is that globally efficient routes can be obtained with relatively little overhead, because these routes need not be maintained at all times. The disadvantage is that reactive routing can entail significant initial delay because the route discovery process is initiated when there is data to send, yet transmission of this data cannot commence until the route discovery process has concluded. The most common protocols for on-demand routing are ad hoc on-demand distance vector routing (AODV) [61] and dynamic source routing (DSR) [62]. Both AODV and DSR aim to minimize the hop count for each source-destination route, but use different techniques to do so. Alternative on-demand routing protocols to AODV and DSR optimize route metrics instead of or in addition to hop count, such as link quality or stability. These alternative protocols are described in [56].

Reactive and proactive routing are combined in a hybrid technique called the zone routing protocol (ZRP). In the ZRP,the network is divided into zones. Each zone establishes routing tables to each of its nodes using a proactive routing technique. When a destination node is outside the zone of the source node, reacting routing is used to find the source-destination route. In ad hoc networks where node communication is primarily local, i.e. within a zone, the ZRP reduces the delay associated with reactive routing within a zone as well as the overhead associated with proactive routing to all nodes in the network since many of these routes would never be used [63]. Hierarchical routing, such as the hierarchical state routing (HSR) protocol, creates clusters of network nodes with one node in each cluster selected as a clusterhead [64]. Within each cluster, routes are established proactively or reactively. When the destination node for a source node lies outside its cluster, the clusterhead communicates with the other clusterheads in the network to determine the appropriate route. Clusters in hierarchical routing are similar to zones in the ZRP; the main difference between hierarchical routing and zone routing is the use of clusterheads to facilitate routing between clusters. Location-based routing, such as the Location-aided routing (LAR) protocol, assumes that the geographical location of nodes in the network are universally known [65]. This location data is typically obtained through the global positioning system (GPS). A route is established by sending packets from a source to the closest node in the direction of the desired destination. Location-based routing generally works poorly in wireless networks because nodes that are close together may not have a strong connection due to fading or shadowing. Moreover, since nodes are generally mobile, locations change constantly, and thus significant overhead is needed to keep an accurate global accounting of mobile locations. Networks where nodes run on batteries often use power aware routing, which minimizes the energy consumption of route discovery as well as packet transmissions over the discovered routes. Power aware routing is discussed in more detail in Chapter 16.4.3.

All routing protocols are based on an underlying network topology: packets can only be routed over links between two nodes of reasonable quality. However, as described earlier, the definition of connectivity between two nodes is somewhat flexible: it depends on the SINR of the link as well as the physical layer design, which determines the required SINR for data to be reliably transmitted over the link. The access layer also plays a role in connectivity, since it dictates the interference between links. Thus, there is significant interaction between the physical, access, and network layers [66]. This interaction was investigated in [42], where it was found that if CSMA with collision avoidance is coupled with a routing protocol that uses low-SINR links, network throughput is significantly reduced. Another interesting result in [42] is that maintaining a single route between any source–destination pair is suboptimal in terms of total network throughput. Multiplexing between multiple routes associated with any given source–destination pair provides an opportunity to change the interference that pair causes to other end-to-end routes, and this diversity can be exploited to increase network throughput.

Routing algorithms can also be optimized for requirements associated with higher-layer protocols, in particular delay and data-rate requirements of the application layer. Such algorithms are referred to as QoS routing. The goal of QoS routing is to find routes through the network that can satisfy the end-to-end delay and data rate requirements specified by the application. Examples of QoS routing and its performance are given in [67, 68, 69].

Most routing protocols use a decode-and-forward strategy at each relay node, where packets received by the relay are decoded to remove errors through error correction and where retransmissions are requested when errors are detected but cannot be corrected. An alternate strategy is amplify-and-forward, where the relay node simply retransmits the packet it has received without attempting to remove errors or detect corrupted packets. This simplifies the relay design, reduces processing energy at the relay, and reduces delay. However, amplify-and-forward does not work well in a wireless setting, since each wireless link introduces noise and fading, which is compounded on each hop of a route. An alternative to these two strategies is *cooperative diversity*, where the diversity associated with spatially distributed users is exploited in forwarding packets [70, 71]. This idea was originally proposed in [72, 73], where multiple transmitters cooperate by repeating detected symbols of the others, thereby forming a repetition code with spatial diversity. These ideas have led to more sophisticated cooperative coding techniques [74] along with forms of cooperative diversity other than coding [75]. The technique of *network*

coding combines different data streams received at intermediate nodes to increase network capacity [76, 77, 78]. Network coding was initially developed to improve efficiency of multicasting in wired networks. Network coding was shown to significantly improve unicast wireless network throughout in [79] by exploiting the fact that packets transmitted from different sources combine "in the air". Wireless network routing protocols that use network coding are described in [80].

16.3 Capacity Limits

The fundamental capacity limits of an ad hoc wireless network – the set of maximum data rates possible between all nodes – is a challenging problem in information theory. For a network of K nodes, each node can communicate with $K-1$ other nodes and so the capacity region has dimension $K(K-1)$ for sending independent information between nodes. The capacity region for sending both independent and common information (multicasting) is much larger. Even for a small number of nodes, the capacity for simple channel configurations (such as the general relay and interference channel) within an ad hoc wireless network remains unsolved [97]. Although rate sums between any two disjoint sets of network nodes can be upper bounded by the corresponding mutual information [97, Thm. 14.10.1], simplifying this formula into a tractable expression for the ad hoc network capacity region is an immensely complex problem. Incorporating delay constraints, feedback, overhead, and outage into the fundamental capacity limit calculation is also challenging, yet needed to provide insights into practical designs [98, 99].

Given that the entire capacity region appears intractable to find, insights can be obtained by focusing on a less ambitious goal. A landmark result by Gupta and Kumar [100] obtained scaling laws for network throughput as the number of nodes in the network K grows asymptotically large. The authors found that the throughput in terms of bits per second for each node in the network decreases with K at a rate between $1/\sqrt{K \log K}$ and $1/\sqrt{K}$. In other words, the per-node rate of the network goes to zero, although the total network throughput (which is equal to the per-node rate multiplied by K) grows at a rate between $\sqrt{K/\log K}$ and $\sqrt{K}$. This surprising result indicates that, even with optimal routing and scheduling, the per-node rate in a large ad hoc wireless network goes to zero. The reason is that intermediate nodes spend much of their resources forwarding packets for other nodes, so few resources are left to send their own data. To some extent this is a pessimistic result: it assumes that nodes choose their destination node at random, whereas in many networks communication between nodes is mostly local. These results were extended in [101] to show that, if mobile nodes can transmit information by physically transporting it close to its desired destination, then node mobility actually increases the per-node rate to a constant; that is, mobility increases network capacity. This increase follows because mobility introduces variation in the network that can be exploited to improve per-user rates. However, in order to exploit the variations due to mobility, significant delay can be incurred. The trade-off between through-put and delay in asymptotically large fixed and mobile networks was characterized in [102, 103]. Similar ideas were applied to finite-size networks and networks with relays in [47, 104].

An alternative approach to scaling laws is to compute achievable rate regions based on suboptimal transmission strategies. This approach was taken in [49] to obtain achievable rate regions based on a time-division strategy associated with all possible rate matrices. The rate matrices describe the set of rates that can be sustained simultaneously by all source–destination pairs at any moment in time. By taking a convex combination of rate matrices at different timeslots, all achievable rates between source–destination pairs under a time-division strategy can be obtained. A rate matrix is a function of the nodes transmitting at that time and the resulting SINR on all links, as well as the transmission strategy. The more capable the transmission strategy, the larger the data rates in a given matrix and the more matrices that are available for use in the time-division scheme. Some of the strategies considered in [49] include variable-rate transmission, single-hop or multihop routing, power control, and successive interference cancellation. The framework can also include the effects of mobility and fading. Figure 16.6 illus-

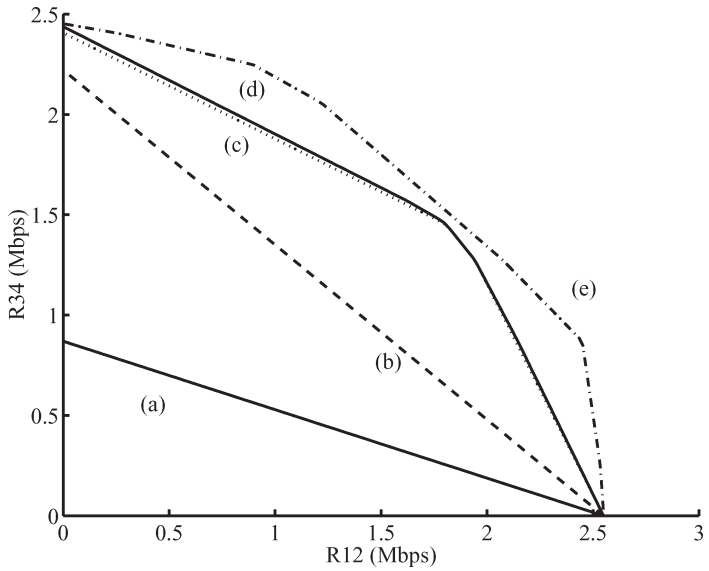


Figure 16.6: Capacity region slice of five-node network along the plane $R_{ij} = 0$, where $\{ij\} \neq \{12\}$, $\{34\}$ and $i \neq j$. (a) Single-hop routing, no spatial reuse. (b) Multihop routing, no spatial reuse. (c) Multihop routing with spatial reuse. (d) Two-level power control added to (c). (e) Successive interference cancellation added to (c).

trates a two-dimensional slice of a rate region for a network of five nodes randomly distributed in a square area. It is assumed that signal propagation between nodes is governed by the simplified path-loss model with path-loss exponent $\gamma = 4$. This two-dimensional slice of the 20-dimensional rate region indicates the rates achievable between two pairs of nodes – from node 1 to 2 and from node 3 to 4 – when all other nodes in the network may be used to help forward traffic between these nodes but do not generate any independent data of their own. The figure assumes variable-rate transmission based on the link SINRs, and it plots the achievable rate region assuming single-hop or multihop routing, spatial reuse, power control, and successive interference cancellation. We see a substantial capacity increase by adding multihop routing, spatial reuse, and interference cancellation. Power control does not provide a significant increase, because adaptive modulation is already being exploited and so adding power control as well does not make much difference – at least for this particular network configuration.

Network capacity regions under different forms of cooperative diversity have also been explored [105, 106, 107, 108, 109]. Since the capacity region of a general ad hoc network is unknown, capacity under cooperation has mainly been characterized by lower bounds based on achievable rate regions or upper bounds based on the rate-sum mutual information bound. Results show that cooperation can lead to substantial gains in capacity, but the advantages of transmitter and/or receiver cooperation – as well as the most advantageous cooperative techniques to use – are highly dependent on network topology and the availability of channel information.

16.4 Energy-Constrained Networks

Many ad hoc wireless network nodes are powered by batteries with a limited lifetime. Thus, it is important to consider the impact of energy constraints in the design of ad hoc wireless networks. Devices with rechargeable batteries must conserve energy in order to maximize time between recharging. In addition, many interesting applications have devices that cannot be recharged–for example, sensors that are imbedded in walls or dropped into a remote region. Such radios must operate for years solely on battery energy and/or energy that can be harvested from the environment. The μ-AMPs and Picoradio projects are aimed at developing radios for these applications

that can operate on less than 100 microwatts and exploit energy harvesting to prolong lifetime [18, 110, 111].

Energy constraints affect the hardware operation, transmit power, and signal processing associated with node operation. The required transmit energy per bit for a given BER target in a noisy channel is minimized by spreading the signal energy over all available time and bandwidth dimensions [112]. However, transmit power is not the only factor in power consumption. The signal processing associated with packet transmission and reception, and even hardware operation in standby mode, consume nonnegligible power as well [28, 113, 114]. This entails interesting energy trade-offs across protocol layers. At the physical layer, many communication techniques that reduce transmit power require a significant amount of signal processing. It is widely assumed that the energy required for this processing is small and continues to decrease with ongoing improvements in hardware technology [28, 115]. However, the results in [113, 114] suggest that these energy costs are still significant. This would indicate that energy-constrained systems must develop energy-efficient processing techniques that minimize power requirements across all levels of the protocol stack and also minimize message passing for network control, since these operations are performed at significant cost to transmitter and receiver energy. Sleep modes for nodes must be similarly optimized, since these modes conserve standby energy but may entail energy costs at other protocol layers due to the associated complications in access and routing. The hardware and operating system design in the node can also be optimized to conserve energy; techniques for this optimization are described in [113, 116]. In fact, energy constraints affect all layers of the protocol stack and hence make cross-layer design even more important if energy-constrained networks are to meet their performance requirements [27, 117, 118, 119]. In this section we describe some of the dominant design considerations for ad hoc wireless networks with energy-constrained nodes.

16.4.1 Modulation and Coding

Modulation and coding choices are typically made based on trade-offs between required transmit power, data rate, BER, and complexity. However, the power consumed within the analog and digital signal processing circuitry can be comparable to the required transmit power for short-range applications. In this case design choices should be based on the total energy consumption, including both the transmit and circuit energy consumption. Modeling circuit energy consumption is quite challenging and depends very much on the exact hardware used [120], which makes it difficult to make broad generalizations regarding trade-offs between circuit and transmit energy. However, the trade-offs certainly exist, especially for short-range applications where transmit energy can be quite low.

Because circuit energy consumption increases with transmission time, minimizing transmission time and then putting nodes to sleep can yield significant energy savings. These ideas were investigated in [110], where it was shown that M-ary modulation may enable energy savings (over binary modulation) for some short-range applications by decreasing the transmission time and shutting down most of the circuitry after transmission. This approach was analyzed for MQAM modulation in [121], which proposed optimal strategies for minimizing the total energy consumption. These ideas were extended in [34] to jointly optimize modulation bandwidth, transmission time, and constellation size for MQAM and MFSK in both AWGN and Rayleigh fading channels. These results indicate that energy consumption is significantly reduced by optimizing transmission time relative to transmission distance: at large distances transmit power dominates, so smaller constellations with larger transmission times are best; but the opposite is true at small transmission distances. As a result, MQAM was slightly more energy efficient than MFSK at short distances because it could transmit over a shorter time duration, but at larger distances MFSK was better owing to the superior energy characteristics of nonlinear amplifiers.

Energy constraints also change the trade-offs inherent to coding. Coding typically reduces the required transmit energy per bit for a given BER target. However, this savings comes at a cost of the processing energy associated with the encoder and decoder. Moreover, some coding schemes (such as block and convolutional codes) encode bits into a codeword that is longer than the original bit sequence; this is sometimes referred to as bandwidth expansion. Although the total transmit energy required for the codeword to achieve a given BER may be less than that required for the uncoded bits, it takes longer for the codeword to be sent, and a longer transmission time consumes more

circuit energy. Joint modulation and coding techniques such as trellis coding do not have bandwidth expansion and hence incur no bandwidth expansion energy penalty. However, their encoder and decoder processing energy must still be taken into account to determine if they yield a net energy savings. The impact of energy constraints on coded MQAM and MFSK was studied in [34]. The results indicate that trellis-coded MQAM provides energy savings at almost all transmission distances of interest (above 1 meter for the hardware parameters considered). However, coding techniques for MFSK are not generally bandwidth efficient, so coding is beneficial for MFSK only at moderate transmissions distances (above 30 meters for the hardware parameters considered).

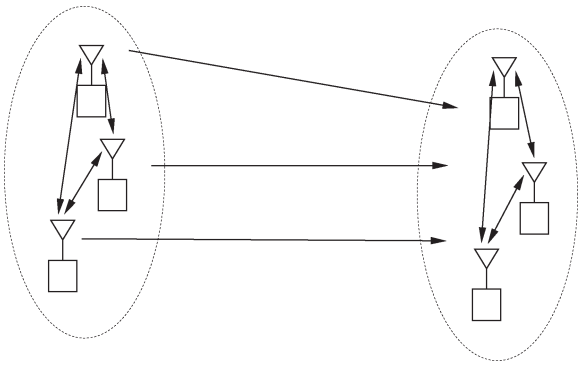


Figure 16.7: Cooperative MIMO.

16.4.2 MIMO and Cooperative MIMO

Multiple-input multiple-output techniques can significantly enhance performance of wireless systems through multiplexing or diversity gain. For a given transmit energy per bit, multiplexing gain provides a higher data rate whereas diversity gain provides a lower BER in fading. However, MIMO systems entail significantly more circuit energy consumption than their single-antenna counterparts, because separate circuitry is required for each antenna signal path and the signal processing associated with MIMO can be highly complex. Thus, it is not clear if MIMO techniques result in performance gain under energy constraints. This question was investigated in [122], where it was found that MIMO does provide energy savings over a single-antenna system for most transmission distances of interest if the constellation size is optimized relative to the distance. The reason is that the MIMO system can support a higher data rate for a given energy per bit, so it transmits the bits more quickly and can then be shut down to save energy.

Many energy-constrained networks consist of very small nodes that cannot support multiple antennas. In this case, nodes that are close together can exchange information to form a multiple-antenna transmitter, and nodes close together on the receive end can cooperate to form a multiple-antenna receiver; see Figure 16.7. As long as the distance between the cooperating nodes is small, the energy associated with their exchange of information is small relative to the energy savings associated with the resulting MIMO system. The energy associated with cooperative MIMO was quantified in [122], where it was shown that this cooperation provides energy savings when the transmit and receive clusters are 10 to 20 times the distance separating the cooperating nodes. When there is less of a difference between the separation of the cooperating nodes and the transmission distance between these clusters, the energy cost required for the local exchange of information exceeds the energy benefits of cooperating. Cooperative MIMO is one form of cooperative diversity. Others were discussed in Section 16.1.4, and these other techniques may provide energy savings comparable to or exceeding those of cooperative MIMO, depending on the network topology.

16.4.3 Access, Routing, and Sleeping

Random access schemes can be made more energy efficient by (i) minimizing collisions and the resulting retransmissions and (ii) optimizing transmit power to the minimum required for successful transmission. One way to reduce collisions is to increase error protection as collisions become more frequent [123]. Alternatively, adaptively minimizing power through probing as part of the random access protocol has been shown to significantly increase energy efficiency [123, 42]. Another method for energy-efficient access is to formulate the distributed access problem using a game-theoretic approach, where energy and delay are costs associated with the game [124]. Several different approaches to energy-efficient access were evaluated in [125]. However, no clear winner emerged because the performance of each protocol is highly dependent on channel characteristics. Delay and fairness constraints can also be incorporated into an energy-efficient access framework, as investigated in [126]. Many of these techniques avoid collisions through a version of TDMA, although setting up channelized access under distributed control can lead to large delays.

If users have long strings of packets or a continuous stream of data, then random access works poorly since most transmissions result in collisions. Hence channels must be assigned to users in a more systematic fashion by transmission scheduling. Energy constraints add a new wrinkle to scheduling optimization. In [112] it was shown that the energy required to send a bit is minimized by transmitting it over all available bandwidth and time dimensions. However, when multiple users wish to access the channel, the system time and bandwidth resources must be shared among all users. More recent work has investigated optimal scheduling algorithms to minimize transmit energy for multiple users sharing a channel [127]. In this work, scheduling was optimized to minimize the transmission energy required by each user subject to a deadline or delay constraint. The energy minimization was based on judiciously varying packet transmission time (and corresponding energy consumption) to meet the delay constraints of the data. This scheme was shown to be significantly more energy efficient than a deterministic schedule with the same deadline constraint.

Energy-constrained networks also require routing protocols that optimize routes relative to energy consumption. If the rate of energy consumption is not evenly distributed across all nodes then some nodes may expire sooner than others, leading to a partitioning of the network. Routing can be optimized to minimize end-to-end energy consumption by applying the standard optimization procedure described in Section 16.1.4, with energy per hop (instead of congestion or delay) as the hop cost [128]. Alternatively, the routes can be computed based on costs associated with the batteries in each node–for example, maximizing the minimum battery lifetime across all nodes in the network [128, 129]. Different cost functions to optimize energy-constrained routing were evaluated via simulation in [128] and were all roughly equivalent. The cost function can also be extended to include the traditional metric of delay along with energy [130]. This method allows the route optimization to trade off between delay and energy consumption through different weighting of their respective contribution to the overall cost function. Note that computation and dissemination of routing tables can entail significant cost: this can be avoided by routing traffic geographically (i.e., in the general direction of its destination), which requires little advance computation [131].

Energy-constrained nodes consume significant power even in standby mode, when they are just passive participants in the network with minimal exchange of data to maintain their network status. The paging industry developed a solution to this problem several decades ago by scheduling "sleep" periods for pagers. The basic idea is that each pager need only listen for transmissions during certain short periods of time. This is a simple solution to implement when a central controller is available, but it is less obvious how to implement such strategies within the framework of distributed network control. Sleep decisions must take into account network connectivity, so it follows that these decisions are local but not autonomous. Mechanisms that support such decisions can be based on neighbor discovery coupled with some means for ordering decisions within the neighborhood. In a given area, the opportunity to sleep should be circulated among the nodes, ensuring that connectivity is not lost through the coincidence of several simultaneous decisions to sleep.

16.4.4 Cross-Layer Design under Energy Constraints

The unique attributes of energy-constrained networks make them prime candidates for cross-layer design. If node batteries cannot be recharged, then each node can transmit only a finite number of bits before it dies, after which time it is no longer available to perform its intended function (e.g. sensing) or to participate in network activities such as routing. Thus, energy must be used judiciously across all layers of the protocol stack in order to prolong network lifetime and meet application requirements.

Energy efficiency at all layers of the protocol stack typically imposes trade-offs between energy consumption, delay, and throughput [132]. However, at any given layer, the optimal operating point on this trade-off curve must be driven by considerations at higher layers. For example, if a node transmits slowly then it conserves transmit energy, but this complicates access for other nodes and increases end-to-end delay. A routing protocol may use a centrally located node for energy-efficient routing, but this will increase congestion and delay on that route and also bum up that node's battery energy quickly, thereby removing it from the network. Ultimately the trade-offs between energy, delay, throughput, and node/network lifetime must be optimized relative to the application requirements. An emergency rescue operation needs on-the-scene information quickly, but typically the network supporting this local information exchange need only last a few hours or days. In contrast, a sensor network embedded into the concrete of a bridge to measure stress and strain must last decades, though the information need only be collected every day or week.

16.4.5 Capacity per Unit Energy

When transmit energy is constrained, it is not possible to transmit any finite number of bits with asymptotically small error probability. This is easy to see intuitively by considering the transmission of a single bit. The only way to ensure that two different values in signal space (representing the two possible bit values) can be decoded with arbitrarily small error is to make their separation arbitrarily large, which requires arbitrarily large energy. Since arbitrarily small error probability is not possible under a hard energy constraint, a different notion of reliable communication is needed for energy-constrained nodes.

A capacity definition for reliable communication under energy constraints was proposed in [133] (see also [112]) as the maximum number of bits per unit energy that can be transmitted so that the probability of error goes to zero with energy. This new notion of capacity per unit energy requires both energy and blocklength to grow asymptotically large for asympotically small error probability. Thus, a finite energy system transmitting at capacity per unit energy does not have an asymptotically small error probability. Insight into this definition for AWGN channels is obtained by examining the minimum energy per bit required to transmit at the normalized Shannon capacity $C_B = C/B$ bps/Hz [112]. Specifically, the received energy per bit equals the ratio of received power to data rate: $E_b = P/R = P/C$ for transmission at rates approaching the Shannon capacity C. Using this expression in the Shannon capacity formula for an AWGN channel yields

$$C = B\log_2\left(1+\frac{P}{N_0B}\right) = B\log_2\left(1+\frac{E_bC}{N_0B}\right). \tag{16.7}$$

Inverting (16.7) yields the energy per bit required to transmit at rates approaching the normalized capacity $C_B = C/B$ as

$$\frac{E_b}{N_0}(C_B) = \frac{2^{C_B}-1}{C_B}. \tag{16.8}$$

As the channel bandwidth B increases, C_B goes to zero, yielding the minimum energy per bit in the wideband limit:

$$\left(\frac{E_b}{N_0}\right)_{\min} = \lim_{C_B \to 0} \frac{2^{C_B} - 1}{C_B} = \ln 2 = -1.59 \text{ dB}. \tag{16.9}$$

It was also shown in [112] that a form of on–off signaling such as pulse position modulation achieves this minimum energy per bit in AWGN. Moreover, results in [134, 135] indicate that the minimum energy per bit for reliable communication in flat fading is also given by (16.9), even when the fading is unknown to the receiver. These results indicate that: (i) in the limit of infinite bandwidth, minimum energy per bit is not affected by fading or receiver knowledge; and (ii) on–off signaling is near optimal for minimum energy communication.

Many energy-constrained wireless systems have large but finite bandwidth, and the results obtained for the limiting case of infinite bandwidth can be misleading for designing such systems (see [135, 136]). In particular, the bandwidth required to operate at an energy per bit close to the minimum (16.9) is very sensitive to the fading distribution and what is known about the fading at the receiver. If fading is known at the receiver then coherent QPSK is the optimal signaling scheme with on–off signaling distinctly suboptimal, but an asymptotic form of on–off signaling is optimal without this receiver knowledge.

The capacity per unit energy for single-user channels has been extended to broadcast and multiple access channels in [133, 138, 137]. These results indicate that, in the wideband limit, TDMA is optimal for both channels because it achieves the minimum energy per bit required for reliable communication. However, in the large but finite bandwidth regime, it was shown in [138] that superposition strategies – such as CDMA coupled with multiuser detection – achieve reliable communication with the same minimum energy per bit and less bandwidth than TDMA.

16.5 Network Types

Because ad hoc wireless networks are generally designed for a particular purpose, they can be broadly categorized into different types that are tailored for specific applications. This section describes some of the most prevalent types of ad hoc wireless networks. The self-configuring nature and lack of infrastructure inherent to these networks make them highly appealing for many applications, even if it results in a significant performance penalty. In particular, the lack of infrastructure in an outdoor data network obviates the need for a large investment to get the network up and running. Deployment costs may then scale with network success. Lack of infrastructure is also highly desirable for military systems, where communication networks must be configured quickly as the need arises, often in remote areas. Other advantages of ad hoc wireless networks include ease of network reconfiguration and reduced maintenance costs. However, these advantages must be balanced against any performance penalty resulting from the multihop routing and distributed control inherent to these networks, which may preclude applications with stringent delay or data rate requirements.

We will focus on the following types of ad hoc wireless networks: data networks, sensor and distributed learning networks, mobile ad hoc networks, and distributed control systems. This list is by no means comprehensive; in fact, the success of ad hoc wireless networks hinges on making them sufficiently flexible so that there can be accidental successes. Therein lies the design dilemma for these networks. If a given network type is designed for maximum flexibility to support many applications (a one-size-fits-all network) then it will be difficult to tailor the network to different application requirements. This will likely result in poor performance for some applications, especially those with high rate requirements or stringent delay constraints. On the other hand, if the network is optimized for a few specific applications then designers must predict in advance what these "killer applications" will be – a risky proposition. Ideally a given type of ad hoc wireless network should be sufficiently flexible to support many different applications while adapting its performance to the applications in operation at any given time. An adaptive cross-layer design can provide this flexibility along with the ability to tailor protocol design to different constraints in the nodes.

16.5.1 Data Networks

Ad hoc wireless data networks connect wireless data devices to each other in order for them to exchange data. In some cases one or more of the nodes is also connected to the backbone Internet. These data networks are generally designed for either indoor or outdoor use. Ad hoc wireless networks designed for indoor use connect devices in homes, offices, or other places where people gather such as stadiums, shopping malls, airports, and train stations. In addition to connecting devices together to exchange data, such networks can stream video to one or more devices, enable smart rooms that sense people and movement and adjust light and heating accordingly, or provide video and motion monitoring with the intelligence to coordinate and interpret data and alert the homeowner, caregiver, or first responder of unusual patterns [10, 11, 12, 13].

Infrastructure-based wireless LANs (Wi-Fi networks) are widely used as indoor data networks, and they generally deliver better performance than ad hoc wireless networks due to their single-hop routing. While ad hoc wireless data networks may provide better coverage indoors than Wi-Fi networks due to mulithop routing between nodes, Wi-Fi networks can also enable multihop routing using a mesh architecturem with multiple access points connected wirelessly. Thus, ad hoc wireless networks are used indoors primary in scenarios where the network must be set up quickly and performance is less of a concern.

Another common type of indoor network supports short-range wireless connections between devices. Such networks are primarily intended to replace inconvenient cabled connections with wireless connections. Thus, the need for cables and the corresponding connectors between cell phones, modems, headsets, computers, printers, projectors,and other such devices is eliminated. The main technology drivers for such networks are low-cost low-power radios with networking capabilities such as Bluetooth [14, 15] and ZigBee [16]. The radios following these standards are integrated into commercial electronic devices to provide networking capabilities between devices. Some common applications include wireless versions of USB and RS232 connectors, PCMCIA cards, set-top boxes, and cell-phone headsets.

Ad hoc wireless networks are needed for outdoor applications where network infrastructure to cover a wide area is too costly or impractical to deploy. In addition, networks that must be built up and torn down quickly (e.g., for military applications or disaster relief) are infeasible without an ad hoc approach. Wireless ad hoc networks that are used outdoors typically require multihop routing, often over many hops when they cover a larger area. The challenge in these networks is to support high data rates in a cost-effective manner and over multiple hops, where the link quality of each hop is different and changes with time. The lack of centralized network control and potential for high-mobility users further complicate this objective. Military programs such as DARPA's GLOMO (global mobile information systems) invested much time and money in building high-speed ad hoc wireless networks for outdoor use to support multimedia, with limited success [7, 8]. Metricom was built as a commercial outdoor ad hoc wireless network in the late 1990s [9]. Although Metricom did deliver fairly high data rates throughout several major metropolitan areas, significant demand never materialized and so the company eventually filed for bankruptcy. In general ad hoc wireless networks outdoors have significantly inferior performance to cellular networks, yet these may be the only communication option for regions the lack cellular or satellite coverage, or for applications where rapid deployment is needed.

16.5.2 Sensor and Distributed Learning Networks

Wireless sensor and distributed learning networks consist of small nodes with sensing, computation, and wireless networking capabilities, and as such these networks represent the convergence of these three important technologies. These networks have enormous potential for both consumer and military applications. Military missions require sensors and other intelligence gathering mechanisms that can be placed close to their intended targets. The potential threat to these mechanisms is therefore quite large, so it follows that the technology used must be highly redundant (or robust) and require as little human intervention as possible. An apparent solution to these constraints

lies in large arrays of passive electromagnetic, optical, chemical, and biological sensors. These can be used to identify and track targets, and they serve also as a first line of detection for various types of attacks. Such networks can also support the movement of unmanned robotic vehicles. For example, optical sensor networks can provide networked navigation, routing vehicles around obstacles while guiding them into position for defense or attack. The design considerations for some industrial applications are quite similar to those for military applications. In particular, sensor arrays can be deployed and used for remote sensing in nuclear power plants, mines, and other industrial venues.

Examples of sensor networks for the home environment include electricity, gas, and water meters that can be read remotely through wireless connections. The broad use of simple metering devices within the home can help monitor and regulate appliances like air conditioners and hot-water heaters that are significant consumers of power and gas. Simple attachments to power plugs can serve as the metering and communication devices for individual appliances. One can imagine a user tracking various types of information on home energy consumption from a single terminal: the home computer. Television usage and content could be tracked or controlled in similar ways. Another important home application is smoke detectors that could not only monitor different parts of the house but also communicate to trace the spread of the fire. Such information (along with house blueprints) could be conveyed to local firefighters before they arrived on the scene. A similar type of array could be used to detect the presence and spread of gas leaks or other toxic fumes.

Sensor arrays also have great potential for use at the sites of large accidents. Consider, for example, the use of remote sensing in rescue operations following the collapse of a building. Sensor arrays could be rapidly deployed at the site of the accident and be used to track heat, natural gas, and toxic substances. Acoustic sensors and localization techniques could be utilized to detect and locate trapped survivors. It may even be possible to prevent such tragedies altogether through the use of sensor arrays. The collapse of bridges, walkways, and balconies, for example, could be predicted in advance using stress and motion sensors built into the structures from the outset. By inserting a large number of these low-cost, low-power sensors directly into the concrete before it is poured, material fatigue could be detected and tracked over time throughout the structure. Such sensors must be robust and self-configuring and would require a long lifetime, commensurate with the lifetime of the structure. Many sensors are deployed with non-rechargeable batteries. The problem of battery lifetime in such sensors may be circumvented by using ultrasmall energy-harvesting radios [18]. Such radios open up new applications for sensor deployment in buildings, homes, and even the human body.

Distributed learning networks are a generalization of sensor networks. The nodes in these networks are generally more powerful than in sensor networks as, in addition to collecting data about their environment and communicating it wirelessly, they are typically equipped with significant memory and computation power to enable machine learning [19]. Decentralized learning allows each node to train on its own localized data which can reduce communication overhead and preserves data privacy. However, as with most decentralized approaches, the performance of distributed learning is typically inferior to centralized learning.

16.5.3 Mobile Ad Hoc Networks

A mobile ad hoc network (MANET) consists of moving objects connected together wirelessly to enable their coordination. Wireless ad hoc networks typically have some moving nodes; what differentiates a MANET is that all the nodes are moving and, in many cases, they are performing a joint task. Examples include vehicular networks, drone networks, and robot networks. The moving objects coordinate either to perform a joint task, such as a robot team, or to perform their individual tasks more efficiently, such as vehicular networks where each vehicle moves towards its desired destination in a more efficient manner through coordination with other vehicles. Node mobility in MANETs allows objects to be moved to a location either to perform their task better or to improve communication in the network. For example, in a robot network, a robot may move far away from the other robots to collect data, or it may move to a centralized location within the team to provide better connectivity among

all the robots. Thus, node movement in a MANET may be optimized to trade off network performance relative to an individual node's capabilities absent mobility constraints. Another unique feature of MANETs is that the movement of the objects may be correlated when the objects are coordinating around a joint task. Fundamental performance limits of MANETs must take into account node mobility, and hence are More details about these characteristics and how they impact MANET design and performance can be found in [20, 21, 22].

16.5.4 Distributed Control Systems

Ad hoc wireless networks also enable distributed control applications, with remote plants, sensors and actuators linked together via wireless communication channels. Such networks allow coordination of unmanned mobile units and feature greatly reduced maintenance and reconfiguration costs in comparison to distributed control systems with wired communication links. Ad hoc wireless networks could be used to support coordinated control of multiple vehicles in an automated highway system, remote control of manufacturing and other industrial processes, and coordination of unmanned airborne vehicles for military applications.

Current distributed control designs provide excellent performance as well as robustness to uncertainty in model parameters. However, these designs are based on closed-loop performance that assumes a centralized architecture, synchronous clocked systems, and fixed topology. Consequently, these systems require that the sensor and actuator signals be delivered to the controller with a small, fixed delay. Ad hoc wireless networks cannot provide any performance guarantees in terms of data rate, delay, or loss characteristics: delays are typically random and packets may be lost. Unfortunately, most distributed controllers are not robust to these types of communication errors, and effects of small random delays can be catastrophic [23, 24]. Thus, distributed controllers must be redesigned for robustness to the random delays and packet losses inherent to wireless networks [25]. Ideally, the ad hoc wireless network can be jointly designed with the controller to deliver the best possible end-to-end performance.

Chapter 16 Problems

1. Consider a signal that must be transmitted 1 km. Suppose the path loss follows the simplified model $P_r = P_t d^{-\gamma}$.

   (a) Find the required transmit power P_t such that the received power is 10 mW for $\gamma = 2$ and $\gamma = 4$.

   (b) Suppose now that a relay node halfway between the transmitter and receiver is used for multihop routing. For $\gamma = 2$ and $\gamma = 4$, if both the transmitter and relay transmit at power P_t, how big must P_t be in order for the relay to receive a signal of 10 mW from the transmitter and the destination to receive a signal of 10 mW from the relay? Find the total power used in the network – that is, the sum of powers at both the transmitter and the relay – for $\gamma = 2$ and $\gamma = 4$.

   (c) Derive a general formula for the total power used in the network with N relays, evenly spaced between the transmitter and receiver, such that each relay and the receiver receive 10 mW of power.

2. Consider an ad hoc wireless network with three users. Users 1 and 2 require a received SINR of 7 dB while user 3 requires an SINR of 10 dB. Assume that all receivers have the same noise power $n_i = 1$ and that there is no processing gain to reduce interference ($\rho = 1$). Assume a matrix of gain values (indexed by the user numbers) of

$$G = \begin{bmatrix} 1 & .06 & .04 \\ .09 & .9 & .126 \\ .064 & .024 & .8 \end{bmatrix}.$$

   (a) Confirm that the equation $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ is equivalent to the SINR constraints of each user.

   (b) Show that a feasible power vector exists for this system such that all users achieve their desired SINRs.

   (c) Find the optimal power vector $\mathbf{P}^*$ such that users achieve their desired SINRs with minimum power.

3. This problem uses the same setup as in the prior problem. Suppose each user starts out with power 50, so the initial power vector $\mathbf{P}(0) = (P_1(0), P_2(0), P_3(0)) = (50, 50, 50)$. Following the recursive formula (16.4) for each user, plot $P_i(k)$ for each of the users $(i = 1, 2, 3)$ over the range $k = 1, \ldots, N$, where N is sufficiently large that the power vector $\mathbf{P}(N)$ is close to its optimal value $\mathbf{P}^*$. Also plot the SINRs of each user for $k = 1, \ldots, N$.

4. Assume an infinite grid of network nodes spaced every 10 meters along a straight line. Assume the simplified path-loss model $P_r = P_t d^{-\gamma}$ and that P_r must be at least 10 mW to establish communication between two nodes.

   (a) For $\gamma = 2$, find $P_{\max}$ such that every node has a neighborhood of $N = 2$ other nodes. What happens if nodes have a peak power constraint less than this $P_{\max}$?

   (b) Find $P_{\max}$ for $\gamma = 2$ and $N = 4$.

   (c) Find $P_{\max}$ for $\gamma = 4$ and $N = 4$.

5. Consider a geographical region of 100 square meters. Suppose that N nodes are randomly distributed in this region according to a uniform distribution and that each node has transmit power sufficient to communicate with any node within a distance of R meters. Compute the average number of nodes $\mathbf{E}[N]$ as a function of radius R, $1 \leq R \leq 20$, required for the network to be fully connected. The average $\mathbf{E}[N]$ should be computed based on 100 samples of the random node placement within the region.

6. Consider a network with multipath routing, so that every packet is sent over N independent routes from its source to its destination. Suppose delay d along each of the multiple end-to-end routes is exponentially distributed with mean D: $p(d) = e^{-d/D}/D,\ d > 0$. Find the probability that all N copies of the packet arrive at the destination with a delay exceeding D as a function of N, and evaluate for $N = 1$ and $N = 5$. Also determine how throughput is affected by multipath routing as a function of N. Qualitatively, describe the throughput/delay trade-off associated with multipath routing.

7. Show that (16.6) is convex in both f_{ij} and C_{ij}.

8. Assume a link with capacity $C = 10$ Mbps. Plot the delay given by (16.5) and by (16.6) for a data flow ranging from 0 to 10 Mbps: $0 \leq f_{ij} < 10$ Mbps.

9. In this problem we consider the difference in the two delay metrics (16.5) and (16.6). Let λ be the ratio of (16.5) to (16.6).

   (a) Find λ as a function of f_{ij}/C_{ij}.

   (b) Using the fact that the flow f_{ij} must be less than C_{ij}, find the range of values that λ can take.

   (c) Find the value of f_{ij}/C_{ij} such that $\lambda > 10$ – that is, where the delay associated with metric (16.5) exceeds that of (16.6) by an order of magnitude.

   (d) Consider a network design where route costs are computed based on either metric (16.5) or (16.6). For which metric will the links be more congested, and why?

10. This problem shows the gains from cross-layer design between the network layer and the application layer. For transmission of applications like video, the application layer will try to use a high-rate encoding scheme to improve the quality. Now consider that the application is real-time. Given a capacity assignment made by the network layer, if the rate of transmission is high then there will be congestion on the link that will delay the packets and hence many packets will not reach the decoder in time, leading to poorer quality. Thus we see a trade-off. A simple model of distortion capturing both these effects can be given as

$$\text{Dist}(R) = D_0 + \frac{\theta}{R - R_0} + \kappa e^{-(C-R)T/L}. \tag{16.10}$$

   The first two terms correspond to distortion at the application layer due to source encoding, and the last term corresponds to distortion due to delayed packets. Let $D_0 = .38,\ R_0 = 18.3$ kbps, $\theta = 2537$, scaling factor $\kappa = 1$, effective packet length $L = 3040$ bits, and playout deadline $T = 350$ ms. Capacity C and transmission rate R are described below.

   (a) If the capacity C takes on the values 45 kbps, 24 kbps, and 60 kbps with probabilities .5, .25, and .25 (respectively), find the optimal rate R that minimizes average distortion Dist (R). Assume full cooperation between the application layer and the network layer: the application layer always knows the instantaneous capacity C associated with the network layer.

   (b) Now consider the case when there is no cross-layer optimization and the application layer encodes at a fixed rate of $R = 22$ kbps at all times. Find the average distortion Dist(R) for the same capacity distribution as given in part (a).

   (c) Comparing parts (a) and (b), find the percentage increase in distortion when cross-layer optimization is not used.

11. Show that the ratio E_b/N_0 as a function of $C_B = C/B$, as given by (16.8), increases with C for B fixed and decreases with B for C fixed. Also show that

$$\lim_{C_B \to 0} \left(\frac{E_b}{N_0} \right) (C_B) = \ln 2.$$

Bibliography

[1] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in packet radio technology," *Proc. IEEE*, pp. 1468–96, November 1978.

[2] L. Kleinrock and J. Silvester, "Optimum transmission radii for packet radio networks or why six is a magic number," *Proc. IEEE Natl. Telecomm. Conf.*, pp. 4.3.1–4.3.5, December 1978.

[3] A. Ephremides, J. E. Wieselthier, and D. J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proc. IEEE*, pp. 56–73, January 1987.

[4] J. Jubin and J. D. Tornow, "The DARPA packet radio network protocols," *Proc. IEEE*, pp. 21–32, January 1987.

[5] M. B. Pursley, "The role of spread spectrum in packet radio networks," *Proc. IEEE*, pp. 116–34, January 1987.

[6] F. A. Tobagi, "Modeling and performance analysis of multihop packet radio networks," *Proc. IEEE*, pp. 135–55, January 1987.

[7] B. Leiner, R. Ruther, and A. Sastry, "Goals and challenges of the DARPA Glomo program (global mobile information systems)," *IEEE Pers. Commun. Mag*., pp. 34–43, December 1996.

[8] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment," *Proc. IEEE Infocom Conf*., pp. 404–13, March 2000.

[9] M. Ritter, R. Friday, and M. Cunningham, "The architecture of Metricom's microcellular data network and details of its implementation as the 2nd and 3rd generation ricochet wide-area mobile data service," *Proc. IEEE Emerg. Technol. Sympos. Broadband Commun*., pp. 143–52, September 2001.

[10] M. N. Huhns, "Networking embedded agents," *IEEE Internet Comput*., pp. 91–3, January/February 1999.

[11] W. W. Gibbs "As we may live," *Scientific American*, pp. 36, 40, November 2000.

[12] K. J. Negus, A. P. Stephens, and J. Lansford, "HomeRF: Wireless networking for the connected home," *IEEE Pers. Commun. Mag*., pp. 20–7, February 2000.

[13] A. Schmidt, "How to build smart appliances," *IEEE Pers. Commun. Mag*., pp. 66–71, August 2001.

[14] J. Haartsen "The Bluetooth radio system," *IEEE Pers. Commun. Mag*., pp. 28–36, February 2000.

[15] J. Haartsen and S. Mattisson, "Bluetooth: A new low-power radio interface providing short-range connectivity," *Proc. IEEE*, pp. 1651–61, October 2000.

[16] I. Poole, “What exactly is . . . ZigBee?” *IEEE Commun. Eng*., pp. 44–5, August/September 2004.

[17] T. Mitchell, “Broad is the way (ultra-wideband technology),” *IEE Review*, pp. 35–9, January 2001.

[18] J. Rabaey, J. Ammer, J. L. da Silva, Jr., and D. Roundy, “PicoRadio supports ad hoc ultra-low power wireless networking,” *IEEE Computer*, pp. 42–8, July 2000.

[19] M. Chen et al., ”Distributed Learning in Wireless Networks: Recent Progress and Future Challenges,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 12, pp. 3579-3605, Dec. 2021

[20] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, “A comprehensive survey on vehicular ad hoc network,”*Journal of network and computer applications*, 37, 380-392. 2014.

[21] I. Bekmezci, O.K. Sahingoz, and Ş. Temel. ”Flying ad-hoc networks (FANETs): A survey.” *Ad Hoc Networks*, Vol. 11, No. 3 (2013): pp. 1254-1270.

[22] P. Basu and J. Redi, ”Movement control algorithms for realization of fault-tolerant ad hoc robot networks,” *IEEE Network*, vol. 18, no. 4, pp. 36-44, July-Aug. 2004.

[23] J. Nilsson, B. Bernhardsson, and B. Wittenmark, “Stochastic analysis and control of real-time systems with random time delays,” *Automatica*, pp. 57–64, January 1998.

[24] X. Liu, S. S. Mahal, A. Goldsmith, and J. K. Hedrick, “Effects of communication delay on string stability in vehicle platoons,” *Proc. IEEE Internat. Conf. Intell. Transp. Syst*., pp. 625–30, August 2001.

[25] S. R. Graham, G. Baliga, and P. R. Kumar, “Issues in the convergence of control with communication and computing: Proliferation, architecture, design, services, and middleware,” *Proc. IEEE Conf. Decision Control*, pp. 1466–71, December 2004.

[26] S. Basagni, D. Turgut, and S. K. Das, “Mobility-adaptive protocols for managing large ad hoc networks,” *Proc. IEEE Internat. Conf. Commun*., pp. 1539–43, June 2001.

[27] N. Bambos, “Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects,” *IEEE Pers. Commun. Mag*., pp. 50–9, June 1998.

[28] K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, “Protocols for self-organization of a wireless sensor network,” *IEEE Pers. Commun. Mag*., pp. 16–27, October 2000.

[29] L. Subramanian and R. H. Katz, “An architecture for building self-configurable systems,” *Proc. Mobile Ad Hoc Network Comput. Workshop*, pp. 63–73, August 2000.

[30] R. Jain, A. Puri, and R. Sengupta, “Geographical routing using partial information for wireless ad hoc networks,” *IEEE Pers. Commun. Mag*., pp. 48–57, February 2001.

[31] L. Zhou and Z. J. Haas, “Securing ad hoc networks,” *IEEE Network*, pp. 24–30, November/December 1999.

[32] R. Karri and P. Mishra, “Modeling energy efficient secure wireless networks using network simulation,” *Proc. IEEE Internat. Conf. Commun*., pp. 61–5, May 2003.

[33] *IEEE Wireless Commun. Mag*., Special Issue on Energy Aware Ad Hoc Wireless Networks (A. J. Goldsmith and S. B. Wicker, Eds.), August 2002.

[34] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, September 2005.

[35] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1992.

[36] V. Kawadia and P. R. Kumar, "Principles and protocols for power control in wireless ad hoc networks," *IEEE J. Sel. Areas Commun.*, pp. 76–88, January 2005.

[37] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Tech.*, pp. 57–62, February 1992.

[38] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Trans. Commun.*, pp. 226–8, February–April 1994.

[39] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Tech.*, pp. 641–6, November 1993.

[40] T. Holliday, N. Bambos, A. J. Goldsmith, and P. Glynn, "Distributed power control for time varying wireless networks: Optimality and convergence," *Proc. Allerton Conf. Commun., Control, Comput.*, pp. 1024–33, October 2003.

[41] N. Bambos, G. Pottie, and S. Chen, "Channel access algorithms with active link protection for wireless communications networks with power control," *IEEE/ACM Trans. Network.*, pp. 583–97, October 2000.

[42] S. Toumpis and A. J. Goldsmith, "Performance, optimization, and cross-layer design of media access protocols for wireless ad hoc networks," *Proc. IEEE Internat. Conf. Commun.*, pp. 2234–40, May 2003.

[43] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, pp. 74–85, January 2004.

[44] S. Kandukuri and N. Bambos, "Power controlled multiple access (PCMA) in wireless communication networks," *Proc. IEEE Infocom Conf.*, pp. 386–95, March 2000.

[45] S. Kallel, "Analysis of memory and incremental redundancy ARQ schemes over a nonstationary channel," *IEEE Trans. Commun.*, pp. 1474–80, September 1992.

[46] V. Rodoplu and T. H. Meng, "Minimum energy mobile wireless networks," *IEEE J. Sel. Areas Commun.*, pp. 1333–44, August 1999.

[47] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. Inform. Theory*, pp. 1877–94, August 2003.

[48] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, pp. 169–81, March 2004.

[49] S. Toumpis and A. Goldsmith, "Capacity regions for ad hoc networks," *IEEE Trans. Wireless Commun.*, pp. 736–48, July 2003.

[50] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," *Proc. IEEE Infocom Conf.*, pp. 1079–88, June 2002.

[51] B. Krishnamachari, S. B. Wicker, and R. Bejar, "Phase transition phenomena in wireless ad hoc networks," *Proc. IEEE Globecom Conf.*, pp. 2921–5, November 2001.

[52] M. Penrose, *Random Geometric Graphs*, Oxford University Press, 2004.

[53] A. Tsirigos and Z. J. Haas, "Multipath routing in the presence of frequent topological changes," *IEEE Commun. Mag*., pp. 132–8, November 2001.

[54] E. M. Royer and Chai-Keong Toh, "A review of current routing protocols for ad hoc mobile wireless networks," *IEEE Personal Communications*, vol. 6, no. 2, pp. 46-55, April 1999.

[55] M.E. M. Campista and M. G. Rubinstein, *Advanced Routing Protocols for Wireless Networks*, Wiley-ISTE, 2014.

[56] A. Boukerche, B. Turgut, N, Aydin, M. Ahmad, L. Bölöni, and D. Turgut, "Routing protocols in ad hoc networks: A survey", *Computer Networks*, Volume 55, Issue 13, pp. 3032-3080, September 2011.

[57] C.-K. Toh, *Ad Hoc Mobile Wireless Networks*: *Protocols and Systems*, Prentice-Hall, Englewood Cliffs, NJ, 2002.

[58] S.-J. Lee and M. Gerla, "Dynamic load-aware routing in ad hoc networks," *Proc. IEEE Internat. Conf. Commun*., pp. 3206–10, June 2001.

[59] L. L. Peterson and B. S. Davie, *Computer Networks – A Systems Approach*, 2nd ed., Morgan-Kaufman, San Mateo, CA, 2000.

[60] C. P. P. Bhagwat, "Highly dynamic destination-sequenced distance vector routing (DSDV) for mobile computers," *Proc. ACM SIGCOMM*, pp. 234–44, September 1994.

[61] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," *Proc. IEEE Workshop Mobile Comput. Syst. Appl*., pp. 90–100, February 1999.

[62] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in T. Imielinsky and H. Korth, Eds., *Mobile Computing*, Kluwer, Dordrecht, 1996.

[63] M. R. Pearlman, Z. J. Haas, and S. I. Mir, "Using routing zones to support route maintenance in ad hoc networks," *Proc. IEEE Wireless Commun. Network Conf*., pp. 1280–4, September 2000.

[64] A. Iwata, C. Chiang, G. Pei, M. Gerla, T. Chen, "Scalable routing strategies for ad hoc wireless networks", *IEEE Journal on Selected Areas in Communications*, 17 (8) (1999) 1369–1379.

[65] Y. Ko, N. Vaidya, "Location-Aided Routing (LAR) in mobile ad hoc networks", *Proceedings of ACM MobiCom*, October 1998, pp. 66– 75.

[66] D. Ayyagari, A. Michail, and A. Ephremides, "A unified approach to scheduling, access control and routing for ad-hoc wireless networks," *Proc. IEEE Veh. Tech. Conf*., pp. 380–4, May 2000.

[67] S. Chen and K. Nahrstedt, "Distributed quality-of-service routing in ad hoc networks," *IEEE J. Sel. Areas Commun*., pp. 1488–1505, August 1999.

[68] C. R. Lin and J.-S. Liu, "QoS routing in ad hoc wireless networks," *IEEE J. Sel. Areas Commun*., pp. 1426–38, August 1999.

[69] M. Mirhakkak, N. Schult, and D. Thomson, "Dynamic bandwidth management and adaptive applications for a variable bandwidth wireless environment," *IEEE J. Sel. Areas Commun*., pp. 1984–97, October 2001.

[70] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag*., pp. 74–80, October 2004.

[71] S. Cui and A. J. Goldsmith, "Energy efficient routing using cooperative MIMO techniques," *Proc. IEEE Internat. Conf. Acous*., *Speech, Signal Proc*., pp. 805–8, March 2005.

[72] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I: System description," *IEEE Trans. Commun*., pp. 1927–38, November 2003.

[73] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun*., pp. 1939–48, November 2003.

[74] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans. Signal Proc*., pp. 362–71, February 2004.

[75] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, pp. 3062–80, December 2004.

[76] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, pp. 1204–16, July 2000.

[77] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Network*., pp. 782–95, October 2003.

[78] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, pp. 371–81, February 2003.

[79] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *IEEE/ACM Transactions on Networking*, Vol. 16, No. 3, June 2008.

[80] S. Sengupta, S. Rayanchu and S. Banerjee, "Network coding-aware routing in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1158-1170, Aug. 2010,

[81] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," *Proc. IEEE Infocom Conf*., pp. 702–11, April 2003.

[82] M. Johansson, L. Xiao, and S. P. Boyd, "Simultaneous routing and power allocation in CDMA wireless data networks," *Proc. IEEE Internat. Conf. Commun*., pp. 51–5, May 2003.

[83] L. Xiao, M. Johansson, and S. P. Boyd, "Optimal routing, link scheduling and power control in multihop wireless networks," *IEEE Trans. Commun*., pp. 1136–44, July 2004.

[84] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, and S.-Y. Kung, "Network planning in wireless ad hoc networks: A cross-layer approach," *IEEE J. Sel. Areas Commun*., pp. 136–50, January 2005.

[85] B. Radunović and J. Y. Boudec, "Rate performance objectives of multihop wireless networks," *IEEE Trans. Mobile Comput*., pp. 334–49, October–December 2004.

[86] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," *Proc. IEEE Infocom Conf*., pp. 477–86, June 2002.

[87] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," *Proc. IEEE*, pp. 974–97, May 1998.

[88] E. Steinbach, N. Farber, and B. Girod, "Adaptive playout for low latency video streaming," *Proc. IEEE Internat. Conf. Image Proc.*, pp. 962–5, October 2001.

[89] D. G. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," *Proc. IEEE Data Compress. Conf.*, p. 569, March 2000.

[90] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," *IEEE Trans. Image Proc.*, pp. 813–26, May 2000.

[91] M. Alasti, K. Sayrafian-Pour, A. Ephremides, and N. Farvardin, "Multiple description coding in networks with congestion problem," *IEEE Trans. Inform. Theory*, pp. 891–902, March 2001.

[92] T. Holliday and A. Goldsmith, "Joint source and channel coding for MIMO systems: Is it better to be robust or quick," *Proc. Joint Workshop Commun. Cod.*, p. 24, October 2004.

[93] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *IEEE Wireless Commun. Mag.*, pp. 3–11, February 2005.

[94] W. Kumwilaisak, Y. T. Hou, Q. Zhang, W. Zhu, C.-C. J. Kuo, and Y. Q. Zhang, "A cross-layer quality-of-service mapping architecture for video delivery in wireless networks," *IEEE J. Sel. Areas Commun.*, pp. 1685–98, December 2003.

[95] T. Yoo, E. Setton, X. Zhu, A. Goldsmith, and B. Girod, "Cross-layer design for video streaming over wireless ad hoc networks," *Proc. IEEE Internat. Workshop Multi. Signal Proc.*, pp. 99–102, September 2004.

[96] W. Yu, K. J. R. Liu, and Z. Safar, "Scalable cross-layer rate allocation for image transmission over heterogeneous wireless networks," *Proc. IEEE Internat. Conf. Acous., Speech, Signal Proc.*, pp. 4.593–4.596, May 2004.

[97] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[98] J. Andrews et al., "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94-101, December 2008.

[99] A. Goldsmith, M. Effros, R. Koetter, M. Médard, A. Ozdaglar and L. Zheng, "Beyond Shannon: the quest for fundamental performance limits of wireless ad hoc networks," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 195-205, May 2011.

[100] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, pp. 388–404, March 2000.

[101] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad-hoc wireless networks," $IEEE/ACM$ *Trans. Network.*, pp. 1877–94, August 2002.

[102] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput–delay trade-off in wireless networks," *Proc. IEEE Infocom Conf.*, pp. 464–75, March 2004.

[103] S. Toumpis, "Large wireless networks under fading, mobility, and delay constraints," *Proc. IEEE Infocom Conf.*, pp. 609–19, March 2004.

[104] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," *Proc. IEEE Infocom Conf.*, pp. 1577–86, June 2002.

[105] A. Host-Madsen, "A new achievable rate for cooperative diversity based on generalized writing on dirty paper," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 317, June 2003.

[106] A. Host-Madsen, "On the achievable rate for receiver cooperation in ad-hoc networks," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 272, June 2004.

[107] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Improved achievable rates for user co-operation and relay channels," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 4, June 2004.

[108] N. Jindal, U. Mitra, and A. Goldsmith, "Capacity of ad-hoc networks with node cooperation," *Proc. IEEE Internat. Sympos. Inform. Theory*, p. 271, June 2004.

[109] C. T. K. Ng and A. J. Goldsmith, "Transmitter cooperation in ad-hoc wireless networks: Does dirty-paper coding beat relaying?" *IEEE Inform. Theory Workshop*, pp. 277–82, October 2004.

[110] A. Chandrakasan, R. Amirtharajah, S. Cho, J. Goodman, G. Konduri, J. Kulik, W. Rabiner, and A. Y. Wang, "Design considerations for distributed microsensor systems," *Proc. IEEE Custom Integrated Circuits Conf.*, pp. 279–86, May 1999.

[111] J. Rabaey, J. Ammer, J. L. da Silva, Jr., and D. Patel, "PicoRadio: Ad-hoc wireless networking of ubiquitous low-energy sensor/monitor nodes," *IEEE Comput. Soc. Workshop on VLSI*, pp. 9–12, April 2000.

[112] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, pp. 1019–30, September 1990.

[113] P. Agrawal, "Energy efficient protocols for wireless systems," *Proc. Internat. Sympos. Pers., Indoor, Mobile Radio Commun.*, pp. 564–9, September 1998.

[114] W. R. Heinzelman, A. Sinha, and A. P. Chandrakasan, "Energy-scalable algorithms and protocols for wireless microsensor networks," *Proc. IEEE Internat. Conf. Acous., Speech, Signal Proc.*, pp. 3722–5, June 2000.

[115] J. M. Kahn, R. H. Katz, and K. S. Pister, "Emerging challenges: Mobile networking for Smart Dust," *J. Commun. Networks*, pp. 188–96, August 2000.

[116] A. Chandrakasan and R. W. Brodersen, *Low Power Digital CMOS Design*, Kluwer, Norwell, MA, 1995.

[117] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Commun. Mag.*, pp. 48–59, August 2002.

[118] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun. Mag.*, pp. 8–27, August 2002.

[119] W. Stark, H. Wang. A. Worthen, S. Lafortune, and D. Teneketzis, "Low-energy wireless communication network design," *IEEE Wireless Commun. Mag.*, pp. 60–72, August 2002.

[120] E. Shih, P. Bahl, and M. Sinclair, "Wake on wireless: An event driven energy saving strategy battery operated devices," *Proc. Internat. Conf. Mobile Comput. Network.*, pp. 160–71, September 2002.

[121] C. Schurgers and M. B. Srivastava, "Energy efficient wireless scheduling: Adaptive loading in time," *Proc. IEEE Wireless Commun. Network Conf.*, pp. 706–11, March 2002.

[122] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO in sensor networks," *IEEE J. Sel. Areas Commun*., pp. 1089–98, August 2004.

[123] M. Zorzi and R. R. Rao, "Energy-constrained error control for wireless channels," *IEEE Pers. Commun. Mag*., pp. 27–33, December 1997.

[124] A. B. MacKenzie and S. B. Wicker, "Selfish users in ALOHA: A game-theoretic approach," *Proc. IEEE Veh. Tech. Conf*., pp. 1354–7, October 2001.

[125] A. Chockalingam and M. Zorzi, "Energy efficiency of media access protocols for mobile data networks," *IEEE Trans. Commun*., pp. 1418–21, November 1998.

[126] J.-H. Youn and B. Bose, "An energy conserving medium access control protocol for multihop packet radio networks," *IEEE Internat. Conf. Comput. Commun. Network*., pp. 470–5, October 2001.

[127] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," $IEEE/ACM$ *Trans. Network*., pp. 487–99, August 2002.

[128] C.-K. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Commun. Mag*., pp. 138–47, June 2001.

[129] J. H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," *Proc. IEEE Infocom Conf*., pp. 609–19, August 2000.

[130] M. A. Youssef, M. F. Younis, and K. A. Arisha, "A constrained shortest-path energy-aware routing algorithm for wireless sensor networks," *Proc. IEEE Wireless Commun. Network Conf*., pp. 794–9, March 2002.

[131] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance," *IEEE Trans. Mobile Comput*., pp. 349–65, October–December 2003.

[132] S. Cui, R. Madan, A. J. Goldsmith, and S. Lall, "Joint routing, MAC, and link layer optimization in sensor networks with energy constraints," *Proc. IEEE Internat. Conf. Commun*., May 2005.

[133] R. G. Gallager, "Energy limited channels: Coding, multiaccess and spread spectrum," *Proc. Inform. Syst. Sci. Conf*., p. 372, March 1988. [See also Technical Report LIDS-P-1714, MIT, November 1987.]

[134] R. S. Kennedy, *Fading Dispersive Communication Channels*, Wiley, New York, 1969.

[135] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, pp. 1319–43, June 2002.

[136] S. Verdú, "Recent results on the capacity of wideband channels in the low-power regime," *IEEE Wireless Commun. Mag*., pp. 40–5, August 2002.

[137] A. Lapidoth, I. E. Telatar, and R. Urbanke, "On wide-band broadcast channels," *IEEE Trans. Inform. Theory*, pp. 3250–8, December 2003.

[138] G. Caire, D. Tuninetti, and S. Verdú, "Suboptimality of TDMA in the low-power regime," *IEEE Trans. Inform. Theory*, pp. 608–20, April 2004.

Appendix A

Representation of Bandpass Signals and Channels

Many signals in communication systems are real bandpass signals with a frequency response that occupies a narrow bandwidth $2B$ centered around a carrier frequency f_c with $2B \ll f_c$, as shown in Figure A.1. Since bandpass signals are real, their frequency response has conjugate symmetry: a bandpass signal $s(t)$ has $|S(f)| = |S(-f)|$ and $\angle S(f) = -\angle S(-f)$. However, bandpass signals are not necessarily conjugate symmetric within the signal bandwidth about the carrier frequency f_c; that is, we may have $|S(f_c + f)| \neq |S(f_c - f)|$ or $\angle S(f_c + f) \neq -\angle S(f_c - f)$ for some $f\colon\ 0 < f \leq B$. This asymmetry in $|S(f)|$ about f_c (i.e., $|S(f_c + f)| \neq |S(f_c - f)|$ for some $f < B$) is illustrated in the figure. Bandpass signals result from modulation of a baseband signal by a carrier, or from filtering a deterministic or random signal with a bandpass filter. The bandwidth $2B$ of a bandpass signal is roughly equal to the range of frequencies around f_c where the signal has nonnegligible amplitude. Bandpass signals are commonly used to model transmitted and received signals in communication systems. These are real signals because the transmitter circuitry can only generate real sinusoids (not complex exponentials), and the channel simply introduces an amplitude and phase change at each frequency of the real transmitted signal.

We begin by representing a bandpass signal $s(t)$ at carrier frequency f_c in the following form:

$$s(t) = s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t), \tag{A.1}$$

where $s_I(t)$ and $s_Q(t)$ are real lowpass (baseband) signals of bandwidth $B \ll f_c$. This is a common representation for bandpass signals or noise. In fact, modulations such as MPSK and MQAM are commonly described using this representation. We call $s_I(t)$ the *in-phase component* of $s(t)$ and $s_Q(t)$ the *quadrature component* of $s(t)$. Define the complex signal $u(t) = s_I(t) + js_Q(t)$, so $s_I(t) = \mathrm{Re}\{u(t)\}$ and $s_Q(t) = \mathrm{Im}\{u(t)\}$. Then $u(t)$ is a complex lowpass signal of bandwidth B. With this definition we see that

$$s(t) = \mathrm{Re}\{u(t)\}\cos(2\pi f_c t) - \mathrm{Im}\{u(t)\}\sin(2\pi f_c t) = \mathrm{Re}\{u(t)e^{j2\pi f_c t}\}. \tag{A.2}$$

The representation on the right-hand side of this equation is called the *complex lowpass representation* of the bandpass signal $s(t)$, and the baseband signal $u(t)$ is called the *equivalent lowpass signal* for $s(t)$ or its *complex envelope*. Note that $U(f)$ is conjugate symmetric about $f = 0$ only if $u(t)$ is real (i.e., if $s_Q(t) = 0$).

Using properties of the Fourier transform, we can show that

$$S(f) = .5[U(f - f_c) + U^*(-f - f_c)]. \tag{A.3}$$

Since $s(t)$ is real, it follows that $S(f)$ is conjugate symmetric about $f = 0$. However, the lowpass signals $U(f)$ and $U^*(f)$ are not necessarily conjugate symmetric about $f = 0$, which leads to an asymmetry of $S(f)$ within the

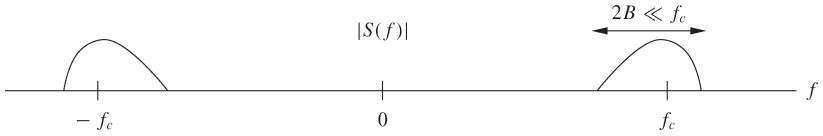


Figure A.1: Bandpass signal $S(f)$.

bandwidth $2B$ about the carrier frequency f_c, as shown in Figure A.1. In fact, $S(f)$ is conjugate symmetric about the carrier frequency within this bandwidth only if $u(t) = s_I(t)$ – that is, if there is no quadrature component in $u(t)$. We will see shortly that this asymmetry affects the response of bandpass channels to bandpass signals.

An alternate representation of the equivalent lowpass signal is

$$u(t) = a(t)e^{j\phi(t)}, \tag{A.4}$$

with envelope

$$a(t) = \sqrt{s_I^2(t) + s_Q^2(t)} \tag{A.5}$$

and phase

$$\phi(t) = \tan^{-1}\left(\frac{s_Q(t)}{s_I(t)}\right). \tag{A.6}$$

With this representation,

$$s(t) = \mathrm{Re}\left\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\right\} = a(t)\cos(2\pi f_c t + \phi(t)). \tag{A.7}$$

Let us now consider a real channel impulse response $h(t)$ with Fourier transform $H(f)$. If $h(t)$ is real then $H^*(-f) = H(f)$. In communication systems we are mainly interested in the channel frequency response $H(f)$ for $|f - f_c| < B$, since only these frequency components of $H(f)$ affect the received signal within the bandwidth of interest. A *bandpass channel* is similar to a bandpass signal: it has a real impulse response $h(t)$ with frequency response $H(f)$ centered at a carrier frequency f_c with a bandwidth of $2B \ll f_c$. To capture the frequency response of $H(f)$ around f_c, we develop an *equivalent lowpass channel* model similar to the equivalent lowpass signal model, as follows. Because the impulse response $h(t)$ corresponding to $H(f)$ is a bandpass signal, it can be written using an equivalent lowpass representation as

$$h(t) = 2\,\mathrm{Re}\{h_l(t)e^{j2\pi f_c t}\}, \tag{A.8}$$

where the extra factor of 2 is to avoid constant factors in the $H(f)$ representation given by (A.9). We call $h_l(t)$ the *equivalent lowpass channel impulse response* for $H(f)$. By (A.2) and (A.3), the representation (A.8) implies that

$$H(f) = H_l(f - f_c) + H_l^*(-f - f_c), \tag{A.9}$$

so $H(f)$ consists of two components: $H_l(f)$ shifted up by f_c, and $H_l^*(f)$ shifted down by f_c. Note that if $H(f)$ is conjugate symmetric about the carrier frequency f_c within the bandwidth $2B$ then $h_l(t)$ will be real and its frequency response $H_l(f)$ will be conjugate symmetric about zero. However, in many wireless channels (e.g., frequency-selective fading channels) $H(f)$ is not conjugate symmetric about f_c, in which case $h_l(t)$ is complex with in-phase component $h_{l,I}(t) = \mathrm{Re}\{h_l(t)\}$ and quadrature component $h_{l,Q}(t) = \mathrm{Im}\{h_l(t)\}$. Note that if $h_l(t)$ is complex then $H_l(f)$ is not conjugate symmetric about zero.

We now use equivalent lowpass signal and channel models to study the output of a bandpass channel with a bandpass signal input. Let $s(t)$ denote the input signal with equivalent lowpass signal $u(t)$. Let $h(t)$ denote the bandpass channel impulse response with equivalent lowpass channel impulse response $h_l(t)$. The transmitted signal $s(t)$ and channel impulse response $h(t)$ are both real, so the channel output $r(t) = s(t) * h(t)$ is also real, with frequency response $R(f) = H(f)S(f)$. Since $S(f)$ is a bandpass signal, $R(f)$ will also be a bandpass signal. Therefore, it has a complex lowpass representation of

$$r(t) = \mathrm{Re}\left\{v(t)e^{j2\pi f_c t}\right\}. \tag{A.10}$$

We now consider the relationship between the equivalent lowpass signals corresponding to the channel input $s(t)$, channel impulse response $h(t)$, and channel output $r(t)$. We can express the frequency response of the channel output as

$$\begin{aligned} R(f) &= H(f)S(f) \\ &= .5[H_l(f-f_c) + H_l^*(-f-f_c)][U(f-f_c) + U^*(-f-f_c)]. \end{aligned} \tag{A.11}$$

For bandpass signals and channels where the bandwidth B of $u(t)$ and $h_l(t)$ is much less than the carrier frequency f_c, we have

$$H_l(f-f_c)U^*(-f-f_c) = 0 \tag{A.12}$$

and

$$H_l^*(-f-f_c)U(f-f_c) = 0. \tag{A.13}$$

Thus,

$$R(f) = .5[H_l(f-f_c)U(f-f_c) + H_l^*(-f-f_c)U^*(-f-f_c)]. \tag{A.14}$$

Given (A.2) and (A.3), (A.10) implies that

$$R(f) = .5[V(f-f_c) + V^*(-f-f_c)]. \tag{A.15}$$

Equating terms at positive and negative frequencies in (A.14) and (A.15), we get that

$$V(f-f_c) = H_l(f-f_c)U(f-f_c) \tag{A.16}$$

and

$$V^*(-f-f_c) = H_l^*(-f-f_c)U^*(-f-f_c) \tag{A.17}$$

or, equivalently, that

$$V(f) = H_l(f)U(f). \tag{A.18}$$

Taking the inverse Fourier transform then yields

$$v(t) = u(t) * h_l(t). \tag{A.19}$$

Thus, we can obtain the equivalent lowpass signal $v(t)$ for the received signal $r(t)$ by taking the convolution of $h_l(t)$ and $u(t)$. The received signal is therefore given by

$$r(t) = \mathrm{Re}\left\{(u(t) * h_l(t))e^{j2\pi f_c t}\right\}. \tag{A.20}$$

Note that $V(f) = H_l(f)U(f)$ is conjugate symmetric about $f = 0$ only if both $U(f)$ and $H_l(f)$ are. In other words, the equivalent lowpass received signal will in general be complex, with nonzero in-phase and quadrature components, if either $u(t)$ or $h_l(t)$ is complex. Moreover, if $u(t) = s_I(t)$ is real (no quadrature component) but the channel impulse response $h_l(t) = h_{l,I}(t) + jh_{l,Q}(t)$ is complex (as with, e.g., frequency-selective fading), then

$$v(t) = s_I(t) * (h_{l,I}(t) + jh_{l,Q}(t)) = s_I(t) * h_{l,I}(t) + js_I(t) * h_{l,Q}(t) \tag{A.21}$$

is complex, so the received signal will have both an in-phase and a quadrature component. More generally, if $u(t) = s_I(t) + js_Q(t)$ and $h_l(t) = h_{l,I}(t) + jh_{l,Q}(t)$ then

$$\begin{aligned} v(t) &= [s_I(t) + js_Q(t)] * [h_{l,I}(t) + jh_{l,Q}(t)] \\ &= [s_I(t) * h_{l,I}(t) - s_Q(t) * h_{l,Q}(t)] + j[s_I(t) * h_{l,Q}(t) + s_Q(t) * h_{l,I}(t)]. \end{aligned} \tag{A.22}$$

Hence the in-phase component of $v(t)$ depends on *both* the in-phase and quadrature components of $u(t)$, and similarly for the quadrature component of $v(t)$. This creates problems in signal detection, since it causes the in-phase and quadrature parts of a modulated signal to interfere with each other in the demodulator.

The main purpose for the equivalent lowpass representation is to analyze bandpass communication systems using the equivalent lowpass models for the transmitted signal, channel impulse response, and received signal. This removes the carrier terms from the analysis – in particular, the dependency of the analysis on the carrier frequency f_c.

Appendix B

Probability Theory, Random Variables, and Random Processes

This appendix provides a brief overview of the main concepts in probability theory, random variables, and random processes that are used throughout the book. More detailed treatments of these broad and deep topics, along with proofs for the properties stated in this appendix, can be found in [1, 2, 3, 4, 5, 6, 7, 8].

B.1 Probability Theory

Probability theory provides a mathematical characterization for random events. Such events are defined on an underlying probability space $(\Omega, \mathcal{E}, p(\cdot))$. The probability space consists of a sample space Ω of possible outcomes for random events; a set of random events $\mathcal{E}$, where each $A \in \mathcal{E}$ is a subset of Ω; and a probability measure $p(\cdot)$ defined on these subsets. Thus, $\mathcal{E}$ is a set of sets, and the probability measure $p(A)$ is defined for every set $A \in \mathcal{E}$. A probability space requires that the set $\mathcal{E}$ be a σ-field. Intuitively, a set $\mathcal{E}$ of sets is a σ-field if it contains all intersections, unions, and complements of its elements.[1] More precisely, $\mathcal{E}$ is a σ-field if: the set of all possible outcomes Ω is one of the sets in $\mathcal{E}$; a set $A \in \mathcal{E}$ implies that $A^c \in \mathcal{E}$; and, for any sets $A_1, A_2, \ldots$ with $A_i \in \mathcal{E}$, we have $\bigcup_{i=1}^{\infty} A_i \in \mathcal{E}$. The set $\mathcal{E}$ must be a σ-field in order for the probability of intersections and unions of random events to be defined. We also require that the probability measure associated with a probability space have the following three fundamental properties.

1. $p(\Omega) = 1$.
2. $0 \leq p(A) \leq 1$ for any event $A \in \mathcal{E}$.
3. If A and B are mutually exclusive (i.e., their intersection is zero), then $p(A \cup B) = p(A) + p(B)$.

Throughout this section, we consider only sets in $\mathcal{E}$, since the probability measure is defined only on these sets.

Several important characteristics of the probability measure $p(\cdot)$ can be derived from its fundamental properties. In particular, $p(A^c) = 1 - p(A)$. Moreover, consider sets $A_1, \ldots, A_n$ where A_i and A_j, $j \neq i$, are disjoint: $A_i \cap A_j = \emptyset$. Then, if $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$, we have $\sum_{i=1}^{n} p(A_i) = 1$. We call the set $\{A_1, \ldots, A_n\}$ with these properties a *partition* of Ω. For two sets A_i and A_j that are not disjoint, $p(A_i \cup A_j) = p(A_i) + p(A_j) - p(A_i \cap A_j)$. This leads to the *union bound*, which states that, for any sets $A_1, \ldots, A_n$,

[1] We use the notation $A \cap B$ to denote the intersection of A and B – that is, all elements in both A and B. The union of A and B, denoted $A \cup B$, is the set of all elements in A or B. The complement of a set $A \subset \Omega$, denoted by A^c, is defined as all elements in Ω that are not in the set A.

$$p(A_1 \cup A_2 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} p(A_i). \tag{B.1}$$

The occurrence of one random event can affect the probability of another random event, since observing one random event indicates which subsets in $\mathcal{E}$ could have contributed to the observed outcome. To capture this effect, we define the probability of event B conditioned on the occurrence of event A as $p(B|A) = p(A \cap B)/p(A)$, assuming $p(A) \neq 0$. This implies that

$$p(A \cap B) = p(A|B)p(B) = p(B|A)p(A). \tag{B.2}$$

The conditional probability $p(B|A) = p(A \cap B)/p(A)$ essentially normalizes the probability of B with respect to the outcomes associated with A, since it is known that A has occurred. We obtain *Bayes' rule* from (B.2) as

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}. \tag{B.3}$$

Independence of events is determined by the probability measure $p(\cdot)$. In particular, events A and B are independent if $p(A \cap B) = p(A)p(B)$. This implies that $p(B|A) = p(B)$ and $p(A|B) = p(A)$.

B.2 Random Variables

Random variables are defined on an underlying probability space $(\Omega, \mathcal{E}, p(\cdot))$. In particular, a random variable X is a function mapping from the sample space Ω to a subset of the real line. If X takes discrete values on the real line it is called a *discrete* random variable, and if it takes continuous values it is called a *continuous* random variable. The *cumulative distribution function* (cdf) of a random variable X is defined as $P_X(x) \triangleq p(X \leq x)$ for some $x \in \mathbb{R}$. The cdf is derived from the underlying probability space as $p(X \leq x) = p(X^{-1}(-\infty, x])$, where $X^{-1}(\cdot)$ is the inverse mapping from the real line to a subset of Ω: $X^{-1}(-\infty, x) = \{\omega \in \Omega : X(\omega) \leq x\}$. Properties of the cdf are based on properties of the underlying probability measure. In particular, the cdf satisfies $0 \leq P_X(x) = p(X^{-1}(-\infty,\ x)) \leq 1$. In addition, the cdf is nondecreasing: $P_X(x_1) \leq P_X(x_2)$ for $x_1 \leq x_2$. This is because $P_X(x_2) = p(X^{-1}(-\infty, x_2)) = p(X^{-1}(-\infty, x_1)) + p(X^{-1}(x_1, x_2)) \geq p(X^{-1}(-\infty, x_1)) = P_X(x_1)$.

The *probability density function* (pdf) of a random variable X is defined as the derivative of its cdf: $p_X(x) \triangleq \frac{d}{dx} P_X(x)$. For X a continuous random variable, $p_X(x)$ is a function over the entire real line; for X a discrete random variable, $p_X(x)$ is a set of delta functions at the possible values of X. The pdf, also referred to as the *probability distribution* or *distribution* of X, determines the probability that X lies in a given range of values:

$$p(x_1 < X \leq x_2) = p(X \leq x_2) - p(X \leq x_1) = P_X(x_2) - P_X(x_1) = \int_{x_1}^{x_2} p_X(x)\, dx. \tag{B.4}$$

Since $P_X(\infty) = 1$ and $P_X(-\infty) = 0$, the pdf integrates to 1:

$$\int_{-\infty}^{\infty} p_X(x)\, dx = 1. \tag{B.5}$$

Note that the subscript X is often omitted from the pdf and cdf when it is clear from the context that these functions characterize the distribution of X. In such cases the pdf is written as $p(x)$ and the cdf as $P(x)$.

The *mean* or *expected value* of a random variable X is its probabilistic average, defined as

$$\mu_X = \mathbf{E}[X] \triangleq \int_{-\infty}^{\infty} x p_X(x)\, dx. \tag{B.6}$$

The expectation operator $\mathbf{E}[\cdot]$ is linear and can also be applied to functions of random variables. In particular, the mean of a function of X is given by

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) p_X(x)\, dx. \tag{B.7}$$

A function of particular interest is the nth *moment* of X,

$$\mathbf{E}[X^n] = \int_{-\infty}^{\infty} x^n p_X(x)\, dx. \tag{B.8}$$

The variance of X is defined in terms of its mean and second moment as

$$\mathrm{Var}[X] = \sigma_X^2 \triangleq \mathbf{E}[(X - \mu_X)^2] = \mathbf{E}[X^2] - \mu_X^2. \tag{B.9}$$

The variance characterizes the average squared difference between X and its mean μ_X. The standard deviation of X, σ_X, is the square root of its variance. From the linearity of the expectation operator it is easily shown that, for any constant c, $\mathbf{E}[cX] = c\mathbf{E}[X]$, $\mathrm{Var}[cX] = c^2\mathrm{Var}[X]$, $\mathbf{E}[X + c] = \mathbf{E}[X] + c$, and $\mathrm{Var}[X + c] = \mathrm{Var}[X]$. Thus, scaling a random variable by a constant scales its mean by the same constant and its variance by the constant squared. Adding a constant to a random variable shifts the mean by the same constant but doesn't affect the variance.

The distribution of a random variable X can be determined from its *characteristic function*, defined as

$$\phi_X(\nu) \triangleq \mathbf{E}[e^{j\nu X}] = \int_{-\infty}^{\infty} p_X(x) e^{j\nu x} dx. \tag{B.10}$$

We see from (B.10) that the characteristic function $\phi_X(\nu)$ of X is the inverse Fourier transform of the distribution $p_X(x)$ evaluated at $\nu/(2\pi)$. Thus we can obtain $p_X(x)$ from $\phi_X(\nu)$ as

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(\nu) e^{-j\nu x} dx. \tag{B.11}$$

This will become significant in finding the distribution for sums of random variables. We can obtain the nth moment of X from $\phi_X(\nu)$ as

$$\mathbf{E}[X^n] = (-j)^n \frac{\partial^n \phi_X(\nu)}{\partial \nu^n}\bigg|_{\nu=0}. \tag{B.12}$$

The *moment generating function* (MGF) of X, defined as $\mathcal{M}_X(\nu) \triangleq \mathbf{E}[e^{\nu X}]$, is similar to the characteristic function. However, it may diverge for some values of ν. If the MGF is finite for ν in the neighborhood around zero, then the nth moment of X is obtained as

$$\mathbf{E}[X^n] = \frac{\partial^n \mathcal{M}_X(\nu)}{\partial \nu^n}\bigg|_{\nu=0}. \tag{B.13}$$

Let X be a random variable and $g(x)$ a function on the real line. Let $Y = g(X)$ define another random variable. Then $P_Y(y) = \int_{x:g(x)\le y} p_X(x)\, dx$. For g monotonically increasing and one-to-one, this becomes $P_Y(y) = \int_{-\infty}^{g^{-1}(y)} p_X(x)\, dx$. For g monotonically decreasing and one-to-one, this becomes $P_Y(y) = \int_{g^{-1}(y)}^{\infty} p_X(x)\, dx$.

We now consider joint random variables. Two random variables must share the same underlying probability space for their joint distribution to be defined. Let X and Y be two random variables defined on the same probability space $(\Omega, \mathcal{E}, p(\cdot))$. Their joint cdf is defined as $P_{XY}(x, y) \triangleq p(X \le x, Y \le y)$. Their joint pdf (distribution) is defined as the derivative of the joint cdf:

$$p_{XY}(x,y) \triangleq \frac{\partial^2 P_{XY}(x,y)}{\partial x \partial y}. \tag{B.14}$$

Thus,

$$P_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p_{XY}(v,w)\,dv\,dw. \tag{B.15}$$

For joint random variables X and Y, we can obtain the distribution of X by integrating the joint distribution with respect to Y:

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)\,dy. \tag{B.16}$$

Similarly,

$$p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x,y)\,dx. \tag{B.17}$$

The distributions $p_X(x)$ and $p_Y(y)$ obtained in this manner are sometimes referred to as the *marginal* distributions relative to the joint distribution $p_{XY}(x,y)$. Note that the joint distribution must integrate to unity:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{XY}(x,y)\,dx\,dy = 1. \tag{B.18}$$

The definitions for joint cdf and joint pdf of two random variables extend in a straightforward manner to any finite number of random variables.

As with random events, observing the value for one random variable can affect the probability of another random variable. We define the conditional distribution of the random variable Y given a realization $X = x$ of random variable X as follows: $p_Y(y|X=x) = p_{XY}(x,y)/p_X(x)$. This implies that $p_{XY}(x,y) = p_Y(y|X=x)p_X(x)$. Independence between two random variables X and Y is determined by their joint distribution. Specifically, X and Y are independent random variables if their joint distribution $p_{XY}(x,y)$ factors into separate distributions for X and Y: $p_{XY}(x,y) = p_X(x)p_Y(y)$. For independent random variables, it is easily shown that $\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\mathbf{E}[g(Y)]$ for any functions $f(x)$ and $g(y)$.

For X and Y joint random variables with joint pdf $p_{XY}(x,y)$, we define their ijth *joint moment* as

$$\mathbf{E}[X^i Y^j] \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^j p_{XY}(x,y)\,dx\,dy. \tag{B.19}$$

The *correlation* of X and Y is defined as $\mathbf{E}[XY]$. The *covariance* of X and Y is defined as $\mathrm{Cov}[XY] \triangleq \mathbf{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbf{E}[XY] - \mu_X\mu_Y$. Note that the covariance and correlation of X and Y are equal if either X or Y has mean zero. The *correlation coefficient* of X and Y is defined in terms of their covariance and standard deviations as $\rho \triangleq \mathrm{Cov}[XY]/(\sigma_X\sigma_Y)$. We say that X and Y are *uncorrelated* if their covariance is zero or, equivalently, if their correlation coefficient is zero. Note that uncorrelated random variables (e.g., X and Y with $\mathrm{Cov}[XY] = \mathbf{E}[XY] - \mu_X\mu_Y = 0$) will have a nonzero correlation ($\mathbf{E}[XY] \neq 0$) if their means are not zero. For random variables $X_1, \ldots, X_n$, we define their *covariance matrix* $\mathbf{\Sigma}$ as an $n \times n$ matrix with ijth element $\Sigma_{ij} = \mathrm{Cov}[X_iX_j]$. In particular, the ith diagonal element of $\mathbf{\Sigma}$ is the variance of X_i: $\mathbf{\Sigma}_{ii} = \mathrm{Var}[X_i]$.

Consider two independent random variables X and Y. Let $Z = X + Y$ define a new random variable on the probability space $(\Omega, \mathcal{E}, p(\cdot))$. We can show directly (or by using characteristic functions) that the distribution of Z is the convolution of the distributions of X and Y: $p_Z(z) = p_X(x) * p_Y(y)$. Equivalently, $\phi_Z(\nu) = \phi_X(\nu)\phi_Y(\nu)$. With this distribution it can be shown that $\mathbf{E}[Z] = \mathbf{E}[X] + \mathbf{E}[Y]$ and that $\mathrm{Var}[Z] = \mathrm{Var}[X] + \mathrm{Var}[Y]$. So, for

sums of independent random variables, the mean of the sum is the sum of the means and the variance of the sum is the sum of the variances.

A distribution that arises frequently in the study of communication systems is the Gaussian distribution. The Gaussian distribution for a random variable X is defined in terms of its mean μ_X and variance σ_X^2 as

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-[(x-\mu_X)^2/(2\sigma_X^2)]}. \tag{B.20}$$

The Gaussian distribution, also called the normal distribution, is denoted as $N(\mu_X, \sigma_X^2)$. Note that the tail of the distribution (i.e., the value of $p_X(x)$ as x moves away from μ_X) decreases exponentially. The cdf $P_X(x) = p(X \leq x)$ for this distribution does not exist in closed form. It is defined in terms of the Gaussian Q-function as

$$P_X(x) = p(X \leq x) = 1 - Q\left(\frac{x - \mu_X}{\sigma_X}\right), \tag{B.21}$$

where the Gaussian Q-function, defined by

$$Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy, \tag{B.22}$$

is the probability that a Gaussian random variable X with mean 0 and variance 1 is larger than x: $Q(x) = p(X \geq x)$ for $X \sim N(0,1)$. The Gaussian Q-function is related to the complementary error function as $Q(x) = .5\,\mathrm{erfc}(x/\sqrt{2})$. These functions are typically calculated using standard computer math packages.

Let $\mathbf{X} = (X_1, \ldots, X_n)$ denote a vector jointly Gaussian random variables. Their joint distribution is given by

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det[\mathbf{\Sigma}]}} \exp[-.5(\mathbf{x} - \mu_\mathbf{X})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mu_\mathbf{X})], \tag{B.23}$$

where $\mu_\mathbf{X} = \mathbf{E}[\mathbf{X}]^T = (\mathbf{E}[X_1], \ldots, \mathbf{E}[X_n])^T$ is the mean of $\mathbf{X}$ and where $\mathbf{\Sigma}$ is the $n \times n$ covariance matrix of $\mathbf{X}$ (i.e., $\mathbf{\Sigma}_{ij} = \mathrm{Cov}[X_i X_j]$). It can be shown from (B.23) that, for jointly Gaussian random variables X and Y, if $\mathrm{Cov}[XY] = 0$ then $p_{XY}(x,y) = p_X(x) p_Y(y)$. In other words, Gaussian random variables that are uncorrelated are independent.

A complex random variable Z is complex Gaussian if $Z = X + jY$ for X and Y jointly Gaussian real random variables. The distribution of Z is then obtained from the joint distribution of X and Y, as given by (B.23) for the vector (X, Y). Similarly, a complex random vector $\mathbf{Z} = (Z_1, \ldots, Z_N) = (X_1 + jY_1, \ldots, X_N + jY_N)$ is complex Gaussian if the random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_N$ are jointly Gaussian real random variables. The distribution of $\mathbf{Z}$ is obtained from joint distribution of these random variables, given by (B.23) for the vector $(X_1, \ldots, X_N, Y_1, \ldots, Y_N)$.

The underlying reason why the Gaussian distribution commonly occurs in communication system models is the *central limit theorem* (CLT), which defines the limiting distribution for the sum of a large number of independent random variables with the same distribution. Specifically, let X_i be independent and identically distributed (i.i.d.) joint random variables. Let $Y_n = \sum_{i=1}^n X_i$ and $Z_n = (Y_n - \mu_{Y_n})/\sigma_{Y_n}$. The CLT states that the distribution of Z_n as n goes to infinity converges to a Gaussian distribution with mean 0 and variance 1: $\lim_{n\to\infty} p_{Z_n}(x) = N(0,1)$. Thus, any random variable equal to the sum of a large number of i.i.d. random components has a distribution that is approximately Gaussian. For example, noise in a radio receiver typically consists of spurious signals generated by the various hardware components, and with a large number of i.i.d. components this noise is accurately modeled as Gauss distributed.

Two other common distributions that arise in communication systems are the uniform distribution and the binomial distribution. A random variable X that is uniformly distributed has pdf $p_X(x) = 1/(b-a)$ for x in

the interval $[a, b]$ and zero otherwise. A random phase θ is commonly modeled as uniformly distributed on the interval $[0, 2\pi]$, which we denote as $\theta \sim \mathcal{U}[0, 2\pi]$. The binomial distribution often arises in coding analysis. Let X_i, $i = 1, \ldots, n$, be discrete random variables that take one of two possible values, 0 or 1. Suppose the X_i are i.i.d. with $p(X_i = 1) = p$ and $p(X_i = 0) = 1 - p$. We call X_i a *Bernoulli* random variable. Let $Y = \sum_{i=1}^{n} X_i$. Then Y is a discrete random variable that takes integer values $k = 0, 1, 2, \ldots$. The distribution of Y is the binomial distribution, given by

$$p(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}, \tag{B.24}$$

where

$$\binom{n}{k} \triangleq \frac{n!}{k!(n-k)!}. \tag{B.25}$$

B.3 Random Processes

A random process $X(t)$ is defined on an underlying probability space $(\Omega, \mathcal{E}, p(\cdot))$. In particular, it is a function mapping from the sample space Ω to a set of real functions $\{x_1(t), x_2(t), \ldots\}$, where each $x_i(t)$ is a possible realization of $X(t)$. Samples of $X(t)$ at times $t_0, t_1, \ldots, t_n$ are joint random variables defined on the underlying probability space. Thus, the joint cdf of samples at times $t_0, t_1, \ldots, t_n$ is given by $P_{X(t_0)X(t_1),\ldots,X(t_n)}(x_0, \ldots, x_n) = p(X(t_0) \leq x_0, X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n)$. The random process $X(t)$ is fully characterized by its joint cdf $P_{X(t_0)X(t_1),\ldots,X(t_n)}(x_0, \ldots, x_n)$ for all possible sets of sample times $\{t_0, t_1, \ldots, t_n\}$.

A random process $X(t)$ is *stationary* if for all T, all n, and all sets of sample times $\{t_0, \ldots, t_n\}$ we have

$$\begin{aligned} p(X(t_0) \leq x_0, X(t_1) \leq & x_1, \ldots, X(t_n) \leq x_n) \\ &= p(X(t_0 + T) \leq x_0, X(t_1 + T) \leq x_1, \ldots, X(t_n + T) \leq x_n). \end{aligned} \tag{B.26}$$

Intuitively, a random process is stationary if time shifts do not affect its probability. Stationarity of a process is often difficult to prove because it requires checking the joint cdf of all possible sets of samples for all possible time shifts. Stationarity of a random process is often inferred from the stationarity of the source generating the process.

The *mean* of a random process is defined as $\mathbf{E}[X(t)]$. Since the mean of a stationary random process is independent of time shifts, it must be constant: $\mathbf{E}[X(t)] = \mathbf{E}[X(t-t)] = \mathbf{E}[X(0)] = \mu_X$. The *autocorrelation* of a random process is defined as $A_X(t, t+\tau) \triangleq \mathbf{E}[X(t)X(t+\tau)]$. The autocorrelation of $X(t)$ is also called its *second moment*. Since the autocorrelation of a stationary process is independent of time shifts, it follows that $A_X(t, t+\tau) = \mathbf{E}[X(t-t)X(t+\tau-t)] = \mathbf{E}[X(0)X(\tau)] \triangleq A_X(\tau)$. Thus, for stationary processes, the autocorrelation depends only on the time difference τ between the samples $X(t)$ and $X(t+\tau)$ and not on the absolute time t. The autocorrelation of a process measures the correlation between samples of the process taken at different times.

Two random processes $X(t)$ and $Y(t)$ defined on the same underlying probability space have joint cdf characterized by

$$\begin{aligned} P_{X(t_0),\ldots,X(t_n),Y(t'_0),\ldots,Y(t'_m)}&(x_0, \ldots, x_n, y_0, \ldots, y_m) \\ &= p(X(t_0) \leq x_0, \ldots, X(t_n) \leq x_n, Y(t'_0) \leq y_0, \ldots, Y(t'_m) \leq y_m) \end{aligned} \tag{B.27}$$

for all possible sets of sample times $\{t_0, \ldots, t_n\}$ and $\{t'_0, \ldots, t'_m\}$. Two random processes $X(t)$ and $Y(t)$ are *independent* if for all such sets we have

$$p_{X(t_0),\ldots,X(t_n),Y(t'_0),\ldots,Y(t'_m)}(X(t_0) \le x_0, \ldots, X(t_n) \le x_n, Y(t'_0) \le y_0, \ldots, Y(t'_m) \le y_m)$$
$$= p_{X(t_0),\ldots,X(t_n)}(X(t_0) \le x_0, \ldots, X(t_n) \le x_n)$$
$$\cdot\, p_{Y(t'_0),\ldots,Y(t'_m)}(Y(t'_0) \le y_0, \ldots, Y(t'_m) \le y_m). \tag{B.28}$$

The cross-correlation between two random processes $X(t)$ and $Y(t)$ is defined as

$$A_{XY}(t, t+\tau) \triangleq \mathbf{E}[X(t)Y(t+\tau)]. \tag{B.29}$$

The two processes are uncorrelated if $\mathbf{E}[X(t)Y(t+\tau)] = \mathbf{E}[X(t)]\mathbf{E}[Y(t+\tau)]$ for all t and τ. As with the autocorrelation, if both $X(t)$ and $Y(t)$ are stationary then the cross-correlation is a function of τ only: $A_{XY}(t, t+\tau) = \mathbf{E}[X(t-t)Y(t+\tau-t)] = \mathbf{E}[X(0)Y(\tau)] \triangleq A_{XY}(\tau)$.

In most analyses of random processes we focus only on the first and second moments. *Wide-sense stationarity* is a notion of stationarity that depends on only the first two moments of a process, and it can be easily verified. Specifically, a process is wide-sense stationary (WSS) if its mean is constant, $\mathbf{E}[X(t)] = \mu_X$, and its autocorrelation depends only on the time difference of the samples, $A_X(t, t+\tau) = \mathbf{E}[X(t)X(t+\tau)] = A_X(\tau)$. Stationary processes are WSS, but WSS processes are not necessarily stationary. For WSS processes, the autocorrelation is a symmetric function of τ, since $A_X(\tau) = \mathbf{E}[X(t)X(t+\tau)] = \mathbf{E}[X(t+\tau)X(t)] = A_X(-\tau)$. Moreover, it can be shown that $A_X(\tau)$ takes its maximum value at $\tau = 0$; that is, $|A_X(\tau)| \le A_X(0) = \mathbf{E}[X^2(t)]$. As with stationary processes, if two processes $X(t)$ and $Y(t)$ are both WSS then their cross-correlation is independent of time shifts and thus depends only on the time difference of the processes: $A_{XY}(t, t+\tau) = \mathbf{E}[X(0)Y(\tau)] = A_{XY}(\tau)$.

The power spectral density (PSD) of a WSS process is defined as the Fourier transform of its autocorrelation function with respect to τ:

$$S_X(f) = \int_{-\infty}^{\infty} A_X(\tau) e^{-j2\pi f\tau} d\tau. \tag{B.30}$$

The autocorrelation can be obtained from the PSD through the inverse transform:

$$A_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f\tau} df. \tag{B.31}$$

The PSD takes its name from the fact that the expected power of a random process $X(t)$ is the integral of its PSD:

$$\mathbf{E}[X^2(t)] = A_X(0) = \int_{-\infty}^{\infty} S_X(f)\, df, \tag{B.32}$$

which follows from (B.31). Similarly, from (B.30) we get that $S_X(0) = \int_{-\infty}^{\infty} A_X(\tau)\, d\tau$. Since $A_X(\tau)$ is real and symmetric, from (B.30) we have that $S_X(f)$ is also symmmetric – that is, $S_X(f) = S_X(-f)$. *White noise* is defined as a zero-mean WSS random process with a PSD that is constant over all frequencies. Thus, a white noise process $X(t)$ has $\mathbf{E}[X(t)] = 0$ and $S_X(f) = N_0/2$ for some constant N_0; this constant is typically referred to as the one-sided white noise PSD. By the inverse Fourier transform, the autocorrelation of white noise is given by $A_X(\tau) = (N_0/2)\delta(\tau)$. In some sense, white noise is the most random of all possible noise processes, since it decorrelates instantaneously.

Random processes are often filtered or modulated, and when the process is WSS the impact of these operations can be characterized in a simple way. In particular, if a WSS process with PSD $S_X(f)$ is passed through a linear time-invariant filter with frequency response $H(f)$, then the filter output is also a WSS process with power spectral density $|H(f)|^2 S_X(f)$. If a WSS process $X(t)$ with PSD $S_X(f)$ is multiplied by a carrier $\cos(2\pi f_c t + \theta)$ with $\theta \sim \mathcal{U}[0, 2\pi]$, the multiplication results in a WSS process $X(t)\cos(2\pi f_c t + \theta)$ with PSD $.25[S_X(f - f_c) + S_X(f + f_c)]$.

A white Gaussian noise process with mean zero and PSD $N_0/2$ can be passed through an ideal bandpass filter centered at frequency f_c with bandwidth $2B$ to create a narrowband noise process $n(t)$. This narrowband process can be represented using the complex lowpass representation as $n(t) = \mathrm{Re}\{n_l(t)e^{j2\pi f_c t}\}$, where $n_l(t) = n_I(t) + jn_Q(t)$ is a complex lowpass Gaussian process with $n_I(t)$ and $n_Q(t)$ real lowpass Gaussian processes. It can be shown that $n_I(t)$ and $n_Q(t)$ are independent, each with mean zero and PSD N_0 (i.e., their PSD is twice that of the original noise process).

Stationarity and WSS are properties of the underlying probability space associated with a random process. We are also often interested in time averages associated with random processes, which can be characterized by different notions of *ergodicity*. A random process $X(t)$ is *ergodic in the mean* if its time-averaged mean, defined as

$$\mu_X^{\mathrm{ta}} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)\, dt, \tag{B.33}$$

is constant for all possible realizations of $X(t)$. In other words, $X(t)$ is ergodic in the mean if $\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x_i(t)\, dt$ equals the same constant μ_X^{ta} for all possible realizations $x_i(t)$ of $X(t)$. Similarly, a random process $X(t)$ is *ergodic in the* nth *moment* if its time-averaged nth moment

$$\mu_{X^n}^{\mathrm{ta}} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X^n(t)\, dt \tag{B.34}$$

is constant for all possible realizations of $X(t)$. We can also define the ergodicity of $X(t)$ relative to its time-averaged autocorrelation

$$A_X^{\mathrm{ta}}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)X(t+\tau)\, dt. \tag{B.35}$$

Specifically, $X(t)$ is *ergodic in autocorrelation* if $\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x_i(t)x_i(t+\tau)\, dt$ equals the same value $A_X^{\mathrm{ta}}(\tau)$ for all possible realizations $x_i(t)$ of $X(t)$. Ergodicity of the autocorrelation in higher-order moments requires that the (nm)th-order time-averaged autocorrelation

$$A_X^{\mathrm{ta}}(n, m, \tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X^n(t)X^m(t+\tau)\, dt \tag{B.36}$$

is constant for all realizations of $X(t)$. A process that is ergodic in all order moments and autocorrelations is called *ergodic*. Ergodicity of a process requires that its time-averaged nth moment and (ij)th autocorrelation, averaged over all time, be constant for all n, i, and j. This implies that the probability associated with an ergodic process is independent of time shifts and thus the process is stationary. In other words, an ergodic process must be stationary. However, a stationary process can be either ergodic or nonergodic. Since an ergodic process is stationary, it follows that

$$\begin{aligned}
\mu_X^{\mathrm{ta}} &= \mathbf{E}[\mu_X^{\mathrm{ta}}] \\
&= \mathbf{E}\left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)\, dt\right] \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{E}[X(t)]\, dt \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mu_X\, dt = \mu_X.
\end{aligned} \tag{B.37}$$

Thus, the time-averaged mean of $X(t)$ equals its probabilistic mean. Similarly,

$$\begin{aligned} A_X^{\text{ta}}(\tau) &= \mathbf{E}[A_X^{\text{ta}}(\tau)] \\ &= \mathbf{E}\left[\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} X(t)X(t+\tau)\,dt\right] \\ &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathbf{E}[X(t)X(t+\tau)]\,dt \\ &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} A_X(\tau)\,dt = A_X(\tau), \end{aligned} \tag{B.38}$$

so the time-averaged autocorrelation of $X(t)$ equals its probabilistic autocorrelation.

B.4 Gaussian Processes

Noise processes in communication systems are commonly modeled as a Gaussian process. A random process $X(t)$ is a Gaussian process if, for all values of T and all functions $g(t)$, the random variable

$$X_g = \int_0^T g(t)X(t)\,dt \tag{B.39}$$

has a Gaussian distribution. A communication receiver typically uses an integrator in signal detection, so this definition implies that, if the channel introduces a Gaussian noise process at the receiver input, then the distribution of the random variable associated with the noise at the output of the integrator will have a Gaussian distribution. The mean of X_g is

$$\mathbf{E}[X_g] = \int_0^T g(t)\mathbf{E}[X(t)]\,dt \tag{B.40}$$

and the variance is

$$\text{Var}[X_g] = \int_0^T \int_0^T g(t)g(s)\mathbf{E}[X(t)X(s)]\,dt\,ds - (\mathbf{E}[X_g])^2. \tag{B.41}$$

If $X(t)$ is WSS then these equations simplify to

$$\mathbf{E}[X_g] = \int_0^T g(t)\mu_X\,dt \tag{B.42}$$

and

$$\text{Var}[X_g] = \int_0^T \int_0^T g(t)g(s)A_X(s-t)\,dt\,ds - (\mathbf{E}[X_g])^2. \tag{B.43}$$

Several important properties of Gaussian random processes can be obtained from this definition. In particular, if a Gaussian random process is input to a linear time-invariant filter, then the filter output is also a Gaussian random process. Moreover, we expect samples $X(t_i), i = 0, 1, \ldots$, of a Gaussian random process to be jointly Gaussian random variables, and indeed this follows from the definition by setting $g(t) = \delta(t - t_i)$ in (B.39). These samples are Gaussian random variables, so if the samples are uncorrelated then they are also independent. In addition, for a WSS Gaussian process, the distribution of X_g in (B.39) depends only on the mean and autocorrelation of the

process $X(t)$. Finally, note that a random process is completely defined by the joint probability of its samples over all sets of sample times. For a Gaussian process, these samples are jointly Gaussian with their joint distribution determined by the mean and autocorrelation of the process. Thus, since the underlying probability of a Gaussian process is completely determined by its mean and autocorrelation, it follows that a WSS Gaussian process is also stationary. Similarly, a Gaussian process that is ergodic in the mean and autocorrelation is an ergodic process.

Bibliography

[1] W. Feller, *An Introduction to Probability Theory and Its Applications*, vols. I and II, Wiley, New York, 1968 and 1971.

[2] R. M. Gray and L. D. Davisson, *Random Processes: A Mathematical Approach for Engineers*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[3] W. B. Davenport, Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1987.

[4] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed., Addison-Wesley, Reading, MA, 1994.

[5] P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, New York, 1995.

[6] R. G. Gallager, *Discrete Stochastic Processes*, Kluwer, Dordrecht, 1996.

[7] H. Stark and J. W. Woods, *Probability and Random Processes with Applications to Signal Processing*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 2001.

[8] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 2002.

Appendix C

Matrix Definitions, Operations, and Properties

This appendix summarizes the definitions, operations, and properties of matrices that are used in the book. More detailed treatments of matrices, along with proofs for the properties stated in this appendix, can be found in [1, 2, 3, 4].

C.1 Matrices and Vectors

An $N \times M$ *matrix* $\mathbf{A}$ is a rectangular array of values with *N* rows and *M* columns, written as

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix}. \tag{C.1}$$

The (*ij*)th element (or entry) of $\mathbf{A}$ (i.e., the element in the ith row and jth column) is written as $\mathbf{A}_{ij}$. In (C.1) we have $\mathbf{A}_{ij} = a_{ij}$. The matrix elements are also called *scalars* to indicate that they are single numbers. An $N \times M$ matrix is called a *square* matrix if $N = M$, a *skinny* matrix if $N > M$, and a *fat* matrix if $N < M$.

The *diagonal elements* of a matrix are the elements along the diagonal line starting from the top left of the matrix (i.e., the elements $\mathbf{A}_{ij}$ with $i = j$). The *trace* of a square $N \times N$ matrix is the sum of its diagonal elements: $\mathrm{Tr}[\mathbf{A}] = \sum_{i=1}^{N} \mathbf{A}_{ii}$. A square matrix is called a *diagonal matrix* if all elements that are not diagonal elements, referred to as the *off-diagonal* elements, are zero: $\mathbf{A}_{ij} = 0$, $j \neq i$. We denote a diagonal matrix with diagonal elements $a_1, \ldots, a_N$ as $\mathrm{diag}[a_1, \ldots, a_N]$. The $N \times N$ identity matrix $\mathbf{I}_N$ is a diagonal matrix with $\mathbf{I}_{ii} = 1$, $i = 1, \ldots, N$; that is, $\mathbf{I}_N = \mathrm{diag}[1, \ldots, 1]$. The subscript N of $\mathbf{I}_N$ is omitted when the size is clear from the context (e.g., from the size requirements for a given operation like matrix multiplication).

A square matrix $\mathbf{A}$ is called *upper triangular* if all its elements below the diagonal are zero (i.e., $\mathbf{A}_{ij} = 0$, $i > j$). A *lower triangular* matrix is a square matrix where all elements above the diagonal are zero (i.e., $\mathbf{A}_{ij} = 0$, $i < j$). Diagonal matrices are both upper triangular and lower triangular.

Matrices can be formed from entries that are themselves matrices, as long as the dimensions are consistent. In particular, if $\mathbf{B}$ is an $N \times M_1$ matrix and $\mathbf{C}$ is an $N \times M_2$ matrix then we can form the $N \times (M_1 + M_2)$ matrix $\mathbf{A} = [\mathbf{B}\ \mathbf{C}]$. The ith row of this matrix is $[\mathbf{A}_{i1} \ldots \mathbf{A}_{i(M_1+M_2)}] = [\mathbf{B}_{i1} \ldots \mathbf{B}_{iM_1} \mathbf{C}_{i1} \ldots \mathbf{C}_{iM_2}]$. The matrix $\mathbf{A}$ formed in this way is also written as $\mathbf{A} = [\mathbf{B}|\mathbf{C}]$. If we also have a $K \times L_1$ matrix $\mathbf{D}$ and a $K \times L_2$ matrix $\mathbf{E}$, then if $M_1 + M_2 = L_1 + L_2$ we can form the $(N + K) \times (M_1 + M_2)$ matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix}. \tag{C.2}$$

The matrices $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$ are called *submatrices* of $\mathbf{A}$. A matrix can be composed of any number of submatrices as long as the sizes are compatible. A submatrix $\mathbf{A}'$ of $\mathbf{A}$ can also be obtained by deleting certain rows and/or columns of $\mathbf{A}$.

A matrix with only one column (i.e., with $M = 1$) is called a *column vector* or just a *vector*. The number of rows of a vector is referred to as its *dimension*. For example, an N-dimensional vector $\mathbf{x}$ is given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}. \tag{C.3}$$

The ith element of vector $\mathbf{x}$ is written as $\mathbf{x}_i$. We call an N-dimensional vector with each element equal to 1 a *ones vector* and denote it by $\mathbf{1}_N$. An N-dimensional vector with one element equal to one and the rest equal to zero is called a *unit vector*. In particular, the ith unit vector $\mathbf{e}^i$ has $\mathbf{e}^i_i = 1$ and $\mathbf{e}^i_j = 0$ for $j \neq i$. A matrix with only one row (i. e., with $N = 1$) is called a *row vector*. The number of columns in a row vector is called its dimension, so an M-dimensional row vector $\mathbf{x}$ is given by $\mathbf{x} = [x_1 \dots x_M]$ with ith element $\mathbf{x}_i = x_i$. The *Euclidean norm* of an N-dimensional row or column vector, also called its *norm*, is defined as

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{N} |\mathbf{x}_i|^2}. \tag{C.4}$$

C.2 Matrix and Vector Operations

If $\mathbf{A}$ is an $N \times M$ matrix then the *transpose* of $\mathbf{A}$, denoted $\mathbf{A}^T$, is the $M \times N$ matrix defined by $\mathbf{A}^T_{ij} = \mathbf{A}_{ji}$:

$$\mathbf{A}^T = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix}^T = \begin{bmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1M} & \cdots & a_{NM} \end{bmatrix}. \tag{C.5}$$

In other words, $\mathbf{A}^T$ is obtained by transposing the rows and columns of $\mathbf{A}$, so the ith row of $\mathbf{A}$ becomes the ith column of $\mathbf{A}^T$. The transpose of a row vector $\mathbf{x} = [\mathbf{x}_1 \dots \mathbf{x}_N]$ yields a vector with the same elements:

$$\mathbf{x}^T = [\mathbf{x}_1 \dots \mathbf{x}_N]^T = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}. \tag{C.6}$$

We therefore often write a column vector $\mathbf{x}$ with elements $\mathbf{x}_i$ as $\mathbf{x} = [\mathbf{x}_1 \dots \mathbf{x}_N]^T$. Similarly, the transpose of an N-dimensional vector $\mathbf{x}$ with ith element $\mathbf{x}_i$ is the row vector $[\mathbf{x}_1 \dots \mathbf{x}_N]$. Note that for $\mathbf{x}$ a row or column vector, $(\mathbf{x}^T)^T = \mathbf{x}$.

The *complex conjugate* $\mathbf{A}^*$ of a matrix $\mathbf{A}$ is obtained by taking the complex conjugate of each element of $\mathbf{A}$:

$$\mathbf{A}^* = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix}^* = \begin{bmatrix} a^*_{11} & \cdots & a^*_{1M} \\ \vdots & \ddots & \vdots \\ a^*_{N1} & \cdots & a^*_{NM} \end{bmatrix}. \tag{C.7}$$

The *Hermitian* of a matrix $\mathbf{A}$, denoted as $\mathbf{A}^H$, is defined as its conjugate transpose: $\mathbf{A}^H = (\mathbf{A}^*)^T$. Note that applying the Hermitian operation twice results in the original matrix: $(\mathbf{A}^H)^H = \mathbf{A}$, so $\mathbf{A}$ is the Hermitian of $\mathbf{A}^H$. A square matrix $\mathbf{A}$ is a *Hermitian matrix* if it equals its Hermitian: $\mathbf{A} = \mathbf{A}^H$. A square matrix $\mathbf{A}$ is a *normal matrix* if $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$. A Hermitian matrix is a normal matrix, since $\mathbf{A}^H = \mathbf{A}$ implies $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$. The complex conjugate and Hermitian operators can also be applied to row or column vectors. In particular, the complex conjugate of a row or column vector $\mathbf{x}$, denoted as $\mathbf{x}^*$, is obtained by taking the complex conjugate of each element of $\mathbf{x}$. The Hermitian of a vector $\mathbf{x}$, denoted as $\mathbf{x}^H$, is its conjugate transpose: $\mathbf{x}^H = (\mathbf{x}^*)^T$.

Two $N \times M$ matrices can be added together to form a new matrix of size $N \times M$. The addition is performed element by element. In other words, if two $N \times M$ matrices $\mathbf{A}$ and $\mathbf{B}$ are added, the resulting $N \times M$ matrix $\mathbf{C} = \mathbf{A} + \mathbf{B}$ has (ij)th element $\mathbf{C}_{ij} = \mathbf{A}_{ij} + \mathbf{B}_{ij}$. Since matrix addition is done element by element, it inherits the commutative and associative properties of addition: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ and $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$. The transpose of a sum of matrices is the sum of the transposes of the individual matrices: $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$. Matrix subtraction is similar: for two $N \times M$ matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{C} = \mathbf{A} - \mathbf{B}$ is an $N \times M$ matrix with (ij)th element $\mathbf{C}_{ij} = \mathbf{A}_{ij} - \mathbf{B}_{ij}$. Two row or column vectors of the same dimension can be added using this definition of matrix addition because such vectors are special cases of matrices. In particular, an N-dimensional column vector $\mathbf{x}$ can be added to another column vector $\mathbf{y}$ of the same dimension to form the new N-dimensional column vector $\mathbf{z} = \mathbf{x} + \mathbf{y}$ with ith element $\mathbf{z}_i = \mathbf{x}_i + \mathbf{y}_i$. Similarly, if $\mathbf{x}$ and $\mathbf{y}$ are row vectors of dimension N then their sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is an N-dimensional row vector with ith element $\mathbf{z}_i = \mathbf{x}_i + \mathbf{y}_i$. However, a row vector of dimension $N > 1$ cannot be added to a column vector of dimension N, since these vectors are matrices of different sizes ($1 \times N$ for the row vector, $N \times 1$ for the column vector). The linear combination of column vectors $\mathbf{x}$ and $\mathbf{y}$ of dimension N yields a new N-dimensional column vector $\mathbf{z} = c\mathbf{x} + d\mathbf{y}$ with ith element $\mathbf{z}_i = c\mathbf{x}_i + d\mathbf{y}_i$, where c and d are arbitrary scalars. Similarly, row vectors $\mathbf{x}$ and $\mathbf{y}$ of dimension N can be linearly combined to form the N-dimensional row vector $\mathbf{z} = c\mathbf{x} + d\mathbf{y}$ with ith element $\mathbf{z}_i = c\mathbf{x}_i + d\mathbf{y}_i$ for arbitrary scalars c and d.

A matrix can be multiplied by a scalar, in which case every element of the matrix is multiplied by the scalar. Specifically, multiplication of the matrix $\mathbf{A}$ by a scalar k results in the matrix $k\mathbf{A}$ given by

$$k\mathbf{A} = k\begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix} = \begin{bmatrix} ka_{11} & \cdots & ka_{1M} \\ \vdots & \ddots & \vdots \\ ka_{N1} & \cdots & ka_{NM} \end{bmatrix}. \tag{C.8}$$

A row vector $\mathbf{x}$ multiplied by scalar k yields $k\mathbf{x} = [k\mathbf{x}_1 \ldots k\mathbf{x}_N]$, and a column vector $\mathbf{x}$ multiplied by scalar k yields $k\mathbf{x} = [k\mathbf{x}_1 \ldots k\mathbf{x}_N]^T$.

Two matrices can be multiplied together provided they have compatible dimensions. In particular, matrices $\mathbf{A}$ and $\mathbf{B}$ can be multiplied if the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$. If $\mathbf{A}$ is an $N \times M$ matrix and $\mathbf{B}$ is an $M \times L$ matrix then their product $\mathbf{C} = \mathbf{A}\mathbf{B}$ is an $N \times L$ matrix with (ij)th element $\mathbf{C}_{ij} = \sum_{k=1}^{M} \mathbf{A}_{ik}\mathbf{B}_{kj}$. Matrix multiplication is not commutative in general (i.e., usually $\mathbf{A}\mathbf{B} \neq \mathbf{B}\mathbf{A}$). In fact, if $\mathbf{A}$ is an $N \times M$ matrix and $\mathbf{B}$ is an $M \times L$ matrix then the product $\mathbf{B}\mathbf{A}$ exists only if $L = N$. In this case $\mathbf{B}\mathbf{A}$ is an $M \times M$ matrix, which may be a different size than the $N \times L$ matrix $\mathbf{A}\mathbf{B}$. Even if $M = L = N$, so that $\mathbf{A}\mathbf{B}$ and $\mathbf{B}\mathbf{A}$ are the same size, they may not be equal. If $\mathbf{A}$ is a square matrix then we can multiply $\mathbf{A}$ by itself. In particular, we define $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$. Similarly $\mathbf{A}^k = \mathbf{A} \cdots \mathbf{A}$ is the product of k copies of $\mathbf{A}$. This implies that $\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l}$. Multiplication of any matrix by the identity matrix of compatible size results in the same matrix; that is, if $\mathbf{A}$ is an $N \times M$ matrix, then $\mathbf{I}_N\mathbf{A} = \mathbf{A}\mathbf{I}_M = \mathbf{A}$. The transpose of a matrix product is the product of the transpose of the individual matrices in reverse order: $(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T$. The product of an $N \times M$ matrix $\mathbf{A}$ and its $M \times N$ Hermitian $\mathbf{A}^H$ is a square matrix. In particular, $\mathbf{A}\mathbf{A}^H$ is an $N \times N$ square matrix while $\mathbf{A}^H\mathbf{A}$ is an $M \times M$ square matrix. The *Frobenius norm* of a matrix $\mathbf{A}$, denoted as $\|\mathbf{A}\|_F$, is defined as $\|\mathbf{A}\|_F = \sqrt{\mathrm{Tr}[\mathbf{A}\mathbf{A}^H]} = \sqrt{\mathrm{Tr}[\mathbf{A}^H\mathbf{A}]} = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M} |\mathbf{A}_{ij}|^2}$. Matrix multiplication is

associative, since $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ as long as the matrix dimensions are compatible for multiplication (hence the parentheses are typically omitted). Matrix multiplication is also distributive: $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ and $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

An M-dimensional vector can be multiplied by a matrix with M columns. Specifically, if $\mathbf{A}$ is an $N \times M$ matrix and $\mathbf{x}$ is an M-dimensional vector (i.e., an $M \times 1$ matrix) then their product yields an N-dimensional vector $\mathbf{y} = \mathbf{Ax}$ with ith element $\mathbf{y}_i = \sum_{k=1}^{M} \mathbf{A}_{ik}\mathbf{x}_k$. Note that a matrix must left-multiply a vector, since the dimensions are not compatible for the product $\mathbf{xA}$. However, if $\mathbf{x}$ is an N-dimensional row vector, then $\mathbf{xA}$ is a compatible multiplication for $\mathbf{A}$ an $N \times M$ matrix and results in the M-dimensional row vector $\mathbf{y} = \mathbf{xA}$ with ith element $\mathbf{y}_i = \sum_{k=1}^{N} \mathbf{x}_k\mathbf{A}_{ki}$. An N-dimensional row vector $\mathbf{x}$ can be multiplied by an N-dimensional vector $\mathbf{y}$, which results in a scalar $z = \mathbf{xy} = \sum_{i=1}^{N} \mathbf{x}_i\mathbf{y}_i$. Note that the transpose of an N-dimensional vector is an N-dimensional row vector. The *inner product* of two N-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{N} \mathbf{x}_i\mathbf{y}_i$.

Given a matrix $\mathbf{A}$, a subset of rows of $\mathbf{A}$ forms a *linearly independent* set if any row in the subset is not equal to a linear combination of the other rows in the subset. Similarly, a subset of columns of $\mathbf{A}$ forms a linearly independent set if any column in the subset is not equal to a linear combination of the other columns in the subset. The *rank* $R_{\mathbf{A}}$ of a matrix $\mathbf{A}$ is equal to the number of rows in the largest subset of linearly independent rows of $\mathbf{A}$, which can be shown to equal the number of columns in the largest subset of linearly independent columns of $\mathbf{A}$. This implies that the rank of an $N \times M$ matrix cannot exceed $\min[N, M]$. An $N \times M$ matrix $\mathbf{A}$ is *full rank* if $R_{\mathbf{A}} = \min[N, M]$.

The *determinant* of a 2×2 matrix $\mathbf{A}$ is defined as $\det[\mathbf{A}] = \mathbf{A}_{11}\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{12}$. For an $N \times N$ matrix $\mathbf{A}$ with $N > 2$, $\det[\mathbf{A}]$ is defined recursively as

$$\det[\mathbf{A}] = \sum_{i=1}^{N} \mathbf{A}_{ij}c_{ij} \tag{C.9}$$

for any j, $1 \leq j \leq N$. Here c_{ij} is the *cofactor* corresponding to the matrix element $\mathbf{A}_{ij}$, defined as

$$c_{ij} = (-1)^{i+j} \det[\mathbf{A}'], \tag{C.10}$$

where $\mathbf{A}'$ is the submatrix of $\mathbf{A}$ obtained by deleting the ith row and jth column of $\mathbf{A}$.

If $\mathbf{A}$ is an $N \times N$ square matrix and if there is another $N \times N$ matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}_N$, then we say that $\mathbf{A}$ is *invertible* or *nonsingular*. We call $\mathbf{B}$ the *inverse* of $\mathbf{A}$, and we denote this inverse as $\mathbf{A}^{-1}$. Thus, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_N$. Moreover, for $\mathbf{A}^{-1}$ defined in this way, we also have that $\mathbf{AA}^{-1} = \mathbf{I}_N$. Only square matrices can be invertible, and the matrix inverse is the same size as the original matrix. A square invertible matrix $\mathbf{U}$ is *unitary* if $\mathbf{UU}^H = \mathbf{I}$, which implies that $\mathbf{U}^H = \mathbf{U}^{-1}$ and thus $\mathbf{U}^H\mathbf{U} = \mathbf{I}$. Not every square matrix is invertible. If a matrix is not invertible, we say it is *singular* or *noninvertible*. The inverse of an inverse matrix is the original matrix: $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. The inverse of the product of two invertible matrices is the product of the inverses in opposite order: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. The kth power of the inverse is $\mathbf{A}^{-k} = (\mathbf{A}^{-1})^k$.

For a diagonal matrix $\mathbf{D} = \mathrm{diag}[d_1, \ldots, d_N]$ with $d_i \neq 0$, $i = 1, \ldots, N$, the inverse exists and is given by $\mathbf{D}^{-1} = \mathrm{diag}[1/d_1, \ldots, 1/d_N]$. For a general 2×2 matrix $\mathbf{A}$ with (ij)th element a_{ij}, its inverse exists if $\det[\mathbf{A}] \neq 0$ and is given by

$$\mathbf{A}^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{\det[\mathbf{A}]} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{C.11}$$

There are more complicated formulas for the inverse of $N \times N$ invertible matrices with $N > 2$. However, matrix inverses are usually obtained using computer math packages.

Matrix inverses are commonly used to solve systems of linear equations. In particular, consider a set of linear equations expressed in matrix form as

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{C.12}$$

If the matrix $\mathbf{A}$ is invertible then, given $\mathbf{y}$, there is a unique vector $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$ that satisfies this system of equations.

C.3 Matrix Decompositions

Given a square matrix $\mathbf{A}$, a scalar value λ for which there exists a nonzero vector $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ is called an *eigenvalue* of $\mathbf{A}$. The vector $\mathbf{x}$ is called the *eigenvector* of $\mathbf{A}$ corresponding to λ. The eigenvalues of a matrix $\mathbf{A}$ are all values of λ that satisfy the *characteristic equation* of $\mathbf{A}$, defined as $\det[\mathbf{A} - \lambda\mathbf{I}] = 0$. The polynomial in λ defined by $\det[\mathbf{A} - \lambda\mathbf{I}]$ is called the *characteristic polynomial* of $\mathbf{A}$, so the eigenvalues of $\mathbf{A}$ are the roots of its characteristic polynomial. The characteristic polynomial of an $N \times N$ matrix has N unique roots $r_1, \ldots, r_N$ with $r_i \neq r_j$ for $i \neq j$, if it is of the form $\det[\mathbf{A} - \lambda\mathbf{I}] = (-1)^N(\lambda - r_1)\cdots(\lambda - r_N)$. When the characteristic polynomial includes a term $(\lambda - r_i)^k, k > 1$, we say that root r_i has *multiplicity k*. For example, if $\det[\mathbf{A} - \lambda\mathbf{I}] = -(\lambda - r_1)^2(\lambda - r_2)^3$ then root r_1 has multiplicity 2 and root r_2 has multiplicity 3. An $N \times N$ matrix has N eigenvalues $\lambda_1, \ldots, \lambda_N$, although they will not all be unique if any of the roots of the characteristic polynomial have multiplicity greater than 1. It can be shown that the determinant of a matrix equals the product of all its eigenvalues (i.e., an eigenvalue r_i with multiplicity k would contribute r_i^k to the product).

The eigenvalues of a Hermitian matrix are always real, although the eigenvectors can be complex. Moreover, if $\mathbf{A}$ is an $N \times N$ normal matrix then it can be written in the following form:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{C.13}$$

where $\mathbf{U}$ is a unitary matrix whose columns are the eigenvectors of $\mathbf{A}$ and $\mathbf{\Lambda} = \mathrm{diag}[\lambda_1, \ldots, \lambda_K, 0, \ldots, 0]$ is an $N \times N$ diagonal matrix whose first K diagonal elements are the nonzero eigenvalues of $\mathbf{A}$. For $\mathbf{A}$ Hermitian, the $\mathbf{\Lambda}$ in (C.13) has only real elements. We say that a matrix $\mathbf{A}$ is *positive definite* if, for all nonzero vectors $\mathbf{x}$, we have $\mathbf{x}^H\mathbf{A}\mathbf{x} > 0$. A Hermitian matrix is positive definite if and only if all its eigenvalues are positive. Similarly, we say the matrix $\mathbf{A}$ is *positive semidefinite* or *nonnegative definite* if, for all nonzero vectors $\mathbf{x}$, $\mathbf{x}^H\mathbf{A}\mathbf{x} \geq 0$. A Hermitian matrix is nonnegative definite if and only if all of its eigenvalues are nonnegative.

Suppose that $\mathbf{A}$ is an $N \times M$ matrix of rank $R_{\mathbf{A}}$. Then there exist an $N \times M$ matrix $\mathbf{\Sigma}$ and two unitary matrices $\mathbf{U}$ and $\mathbf{V}$ of size $N \times N$ and $M \times M$ (respectively) such that

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \tag{C.14}$$

We call the columns of $\mathbf{V}$ the *right singular vectors* of $\mathbf{A}$ and the columns of $\mathbf{U}$ the *left singular vectors* of $\mathbf{A}$. The matrix $\mathbf{\Sigma}$ has a special form: all elements that are not diagonal elements are zero, so

$$\mathbf{\Sigma}_{N\times M} = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_M \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \tag{C.15}$$

for $N \geq M$ and

$$\mathbf{\Sigma}_{N\times M} = \begin{bmatrix} \sigma_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N & 0 & \cdots & 0 \end{bmatrix} \tag{C.16}$$

for $N < M$, where $\sigma_i = \sqrt{\lambda_i}$ for λ_i the ith eigenvalue of $\mathbf{A}\mathbf{A}^H$. The values of σ_i are called the *singular values* of $\mathbf{A}$, and $R_\mathbf{A}$ of these singular values are nonzero, where $R_\mathbf{A}$ is the rank of $\mathbf{A}$. The decomposition (C.14) is called the *singular value decomposition* (SVD) of $\mathbf{A}$. The singular values of a matrix are always nonnegative.

Let $\mathbf{A}$ be an $N \times M$ matrix whose ith column we denote as $\mathbf{A}_i$. Treating each column as a submatrix, we can write $\mathbf{A} = [\mathbf{A}_1\ \mathbf{A}_2 \cdots \mathbf{A}_M]$. The *vectorization* of the matrix $\mathbf{A}$, denoted as vec($\mathbf{A}$), is defined as the (NM)-dimensional vector that results from stacking the columns $\mathbf{A}_i$, $i = 1, \ldots, M$, of matrix $\mathbf{A}$ on top of each other to form a vector:

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_M \end{bmatrix} = [\mathbf{A}_{11}\ \mathbf{A}_{21} \ldots \mathbf{A}_{N1}\ \mathbf{A}_{12} \ldots \mathbf{A}_{N2} \ldots \mathbf{A}_{1M} \cdots \mathbf{A}_{NM}]^T. \tag{C.17}$$

Let $\mathbf{A}$ be an $N \times M$ matrix and $\mathbf{B}$ an $L \times K$ matrix. The *Kronecker product* of $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{A} \otimes \mathbf{B}$, is an $NL \times MK$ matrix defined by

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B} & \cdots & \mathbf{A}_{1M}\mathbf{B} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{N1}\mathbf{B} & \cdots & \mathbf{A}_{NM}\mathbf{B} \end{bmatrix}. \tag{C.18}$$

Bibliography

[1] B. Nobel and J. W. Daniel, *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[2] G. Strang, *Linear Algebra and Its Applications*, 2nd ed., Academic Press, New York, 1980.

[3] R. A. Hom and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[4] R. A. Hom and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

Appendix D

Summary of Wireless Standards

This appendix summarizes the technical details associated with the two most prevalent wireless systems in operation today: cellular phones and wireless LANs. It also summarizes the specifications for three short range wireless network standards that support a broad range of applications.

D.1 Cellular Phone Standards

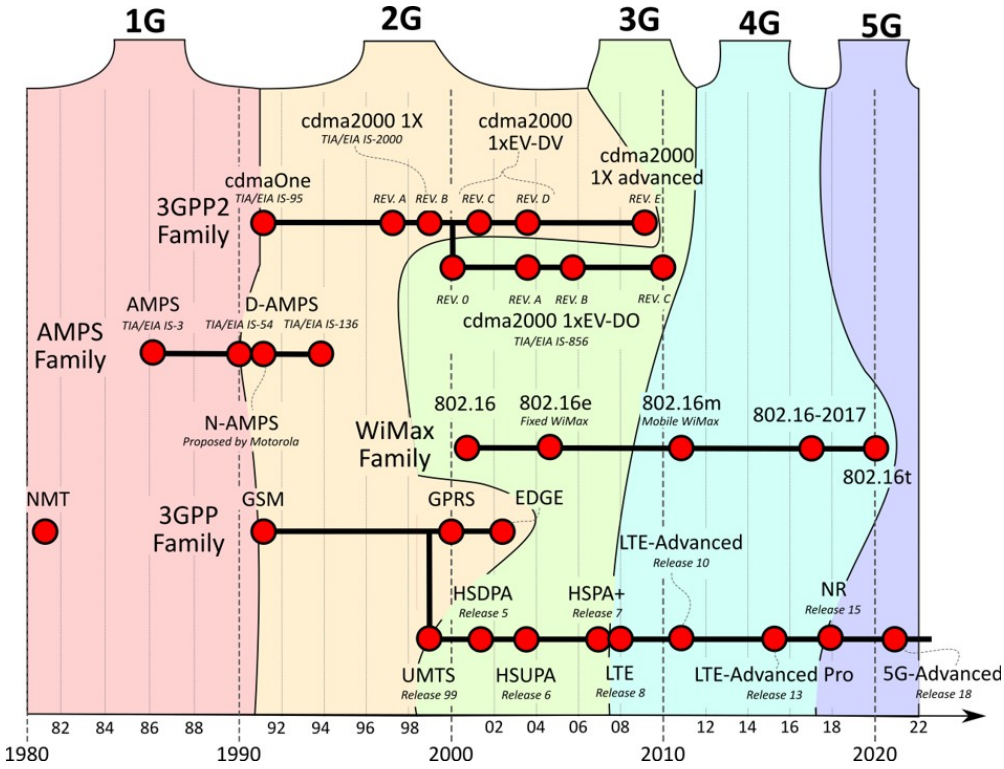


Figure D.1: Timeline of Cellular Network Standards.

In this section we summarize cellular phone standards, which began with the first generation (1G) analog systems deployed in the 1980s and has evolved approximately every decade to a new standard [1, 2, 3]. A timeline for these different standards is shown in the Fig D.1, taken from [4]. As this timeline indicates, in between new

generations, the standard has evolved throughout the decade via continuous new standard releases. The figure also indicates that up until 5G, there were competing standards for each generation, although LTE and LTE Advanced standards developed by the third-generation partnership project (3GPP) became the dominant standard for 4G, with mobile Wimax (802.16e) relegated to niche applications and geographies. The fifth generation of cellular networks around the world adopted the 3GPP 5G standard, finally uniting all cellular systems worldwide around the same standard.

A summary of the key technical features in each generation of cellular standard is show in Table D.1 below. The biggest technological shifts happened when the standard moved from analog to digital (1G to 2G), converged on CDMA as the underlying physical layer technology for 3G, and then moved to OFDM and MIMO as the underlying physical layer technologies coupled with a packet switched network architecture for 4G. The major change in moving from 4G to 5G as the introduction of network slicing, while the physical layer changes were incremental. The proposed physical and network layer changes from 5G to 6G are still under development with major changes unexpected, indicative that these technologies are maturing.

Feature	1G (analog)	2G	2.5G	3G	4G	4G+ (LTE Advanced)	5G	6G (Anticipated)
One Global Standard	No	No	No	Yes	Yes	Yes	Yes	Yes
Peak Data Rate	N/A	42-270 Kbps	57-384 Kbps	42 Mbps	1 Gbps	3 Gbps	20 Gbps	100 Gbps
Latency	500 ms	500 ms	500 ms	100-500 ms	20-30 ms	10-20 ms	1-4 ms	< 1ms
Spectral Band Locations	400 MHz, 800 MHz, 900 MHz	800 MHz, 900 MHz, 1.4 GHz, 1.5 GHz	800 MHz, 900 MHz, 1.4 GHz, 1.5 GHz, 1.9 GHz	800 MHz 2.1 GHz	600 MHz, 800 MHz, 2.5 GHz	600 MHz, 800 MHz, 2.5 GHz, 6 GHz	600 MHz, 800 MHz, 2.5 GHz, 6 GHz, 60 GHz	600 MHz, 800 MHz, 2.5 GHz, 6 GHz, 60 GHz, 90-300 GHz
Channel Bandwidth	10-30 KHz	25 KHz-1.25 MHz	25 KHz-1.25 MHz	1.25-5 MHz	Up to 20 MHz	Up to 100 MHz	Up to 100 MHz (600 MHz-6 GHz) , up to 400 MHz (60 GHz)	Up to 100 MHz (600 MHz-6 GHz) , up to 400 MHz (60 GHz). Up to 1 GHz (90-300 GHz)
Downlink Speed	N/A	42-270 Kbps	57-384 Kbps	3-7 Mbps	10-50 Mbps	100-150 Mbps	100 Mbps – 10 Gbps	10-100 Gbps
Uplink Speed	N/A	42-270 Kbps	57-384 Kbps	1 Mbps	10 Mbps	50 Mbps	100 Mbps – 10 Gbps	10-100 Gbps
Main Applications	Voice	Voice	Voice, Messaging	Voice, Messaging (SMS), Low Speed Data	Low-Res Video, Gaming, Data Transfer	Video, Gaming, High Speed Data	Video, Gaming, High Speed Data	Video, Gaming, High Speed Data, AR/VR, IoT
Network Architecture	Circuit Switched	Circuit Switched	Circuit Switched	Circuit Switched	Packet Switched	Packet Switched	Packet Switched, Network Slicing	Packet Switched, Network Slicing
Multiple Access	FDMA	TDMA with slow FH or CDMA	TDMA with slow FH or CDMA	CDMA	OFDMA	OFDMA	OFDMA	OFDMA
Modulation	FM	GMSK, BPSK, QPSK, π/4 D-QPSK	GMSK, BPSK, QPSK, π/4-D-QPSK	BPSK, QPSK, MPSK	Adaptive MQAM	Adaptive MQAM	Adaptive MQAM	Adaptive MQAM
Coding	N/A	Convolutional, Parity Check	Convolutional, Parity Check	Convolutional, Turbo	Turbo, LDPC	Turbo, LDPC	Turbo, LDPC, Polar	Turbo, LDPC, Polar
ISI Mitigation	Equalizer	Equalizer or CDMA + RAKE Receiver	Equalizer or CDMA + RAKE Receiver	CDMA + RAKE Receiver	OFDM	OFDM	OFDM	OFDM
Multiple Antennas (MIMO)	None	None	None	None	2x2	4 at the BS, Up to 4 at the UE	Up to 128 at the BS, Up to 8 at the UE	Up to 1028 or more at the BS, up to 8 at the UE

Table D.1: Key Features of the 1G to 6G Cellular Standards

The remainder of this section will describe each generation of cellular standard, the "standards wars" and which standard emerged as dominant, and how each generation of the standard differed from the previous generation.

First-Generation Analog Systems

The main characteristics of first-generation (1G) analog cellular standard are summarized in Table D.2. Systems based on these standards were widely deployed in the 1980s. The best known standard is the Advanced Mobile Phone Service (AMPS), developed by Bell Labs in the 1970s and first used commercially in the United States in 1983. After its U.S. deployment, many other countries adopted AMPS as well. This system has a nar-

System	AMPS	TACS	NMT	NTT	C-450	RC2000
Uplink Frequencies (MHz)	824-849	890-915	453-458, 890-915	925-940*	450-455.74	414.8-418**
Downlink Frequencies (MHz)	869-894	933-960	463-468, 935-960	870-885	460-465.74	424.8-428
Channel Bandwidth (KHz)	30	25	25/12.5	25	10	12.5
Number of Channels	832	1000	180, 1999	600	573	256
Modulation/Multiple Access	FM/FDMA					
* NTT operated in several other frequency bands around 900 MHz ** RC2000 operated in several other frequency bands around 200 MHz						

Table D.2: First-generation analog cellular phone standards

rowband version, narrowband AMPS (N-AMPS), with voice channels that are one third the bandwidth of regular AMPS. Japan deployed the first commercial cellular phone system in 1979 with the NTT (MCS-L1) standard based on AMPS, but at a higher frequency and with voice channels of slightly lower bandwidth. Europe also developed a similar standard to AMPS called the Total Access Communication System (TACS), which operates at a higher frequency and with smaller bandwidth channels than AMPS. It was deployed in the United Kingdom and in other European countries as well as outside Europe. The frequency range for TACS was extended in the U.K. to obtain more channels, leading to a variation called ETACS. A variation of the TACS system called JTACS was deployed in metropolitan areas of Japan in 1989 to provide higher capacity than the NTT system. JTACS operates at a slightly higher frequency than TACS or ETACS, and it has a bandwidth-efficient version (called NTACS) where voice channels occupy half the bandwidth of the channels in JTACS. In addition to TACS, countries in Europe had different incompatible standards at different frequencies for analog cellular, including the Nordic Mobile Telephone (NMT) standard in Scandinavia, the Radiocom 2000 (RC2000) standard in France, and the C-450 standard in Germany and Portugal. The incompatibilities made it impossible to roam between European countries with a single analog phone, which motivated the need for one unified cellular standard and frequency allocation throughout Europe.

Second-Generation Digital Systems

System	GSM	IS-136	IS-95 (cdmaOne)	PDC
Uplink Frequencies (MHz)	890-915	824-849	824-849	940-956, 1429-1453
Downlink Frequencies (MHz)	935-960	869-894	869-894	810-826, 1477-1501
Channel Bandwidth (KHz)	200	30	1250	25
Number of Channels	1000	832 (3 users/channel)	2500	1600 (3 users/channel)
Channel Data Rate (Kbps)	270.833	48.6	1.2288 Mchips/sec	42
Compressed Speech Rate (Kbps)	13	7.95	1.2-9.6 (Adaptive)	6.7
Multiple Access	TDMA with Slow FH	TDMA	CDMA with Soft Handoff	TDMA
Modulation	GMSK	$\pi/4$-D-QPSK	BPSK, QPSK	$\pi/4$-D-QPSK
Coding	Convolutional (Rate 1/2)	Convolutional (Rate 1/2)	Convolutional (Rate 1/2 Downlink, 1/3 Uplink)	Convolutional (Rate 1/2)
ISI Mitigation	Equalizer	Equalizer	RAKE Receiver	Equalizer

Table D.3: Second-generation (digital) cellular phone standards

Next we consider second-generation (2G) digital cellular phone standards, whose main characteristics are summarized in Table D.3. These systems were mostly deployed in the early 1990s. As a result of the incompatibilities in the first-generation analog systems, in 1982 the Groupe Spécial Mobile (GSM) was formed to develop a uniform digital cellular standard for all of Europe. The TACS spectrum in the 900-MHz band was allocated for GSM operation across Europe in order to facilitate roaming between countries. In 1989 the GSM specification was finalized and the system was launched in 1991, although availability was limited until 1992. The GSM standard uses TDMA combined with slow frequency hopping to combat out-of-cell interference. Convolutional coding and parity-check codes along with interleaving are used for error detection and correction. The standard also includes an equalizer to compensate for frequency-selective fading. At its peak the GSM standard was used in about 66%

of the world's cell phones, with more than 470 GSM operators in 172 countries supporting over a billion users. As the GSM standard became more global, the meaning of the acronym was changed to the Global System for Mobile Communications.

Although Europe got an early jump on developing 2G digital systems, the United States was not far behind. In 1992 the IS-54 digital cellular standard was finalized, with commercial deployment beginning in 1994. This standard uses the same channel spacing (30 kHz) as AMPS to facilitate the analog-to-digital transition for wireless operators, along with a TDMA multiple access scheme to improve handoff and control signaling as compared with analog FDMA. The IS-54 standard (also called the North American Digital Cellular standard) was improved over time, and these improvements evolved into the IS-136 standard, which subsumed the original standard. Similar to the GSM standard, the IS-136 standard uses parity-check codes, convolutional codes, interleaving, and equalization.

A competing standard for 2G systems based on CDMA was proposed by Qualcomm in the early 1990s. The standard, called IS-95 or IS-95a, was finalized in 1993 and deployed commercially under the name cdmaOne in 1995. Like IS-136, IS-95 was designed to be compatible with AMPS so that the two systems could coexist in the same frequency band. In CDMA all users are superimposed on top of each other with spreading codes that can separate out the users at the receiver. Thus, channel data rate does not apply to just one user, as in TDMA systems. The channel chip rate is 1.2288 Mchips per second for a total spreading factor of 128 for both the uplink and downlink. The spreading process in IS-95 is different for the downlink (DL) and the uplink (UL), with spreading on both links accomplished through a combination of spread-spectrum modulation and coding. On the downlink, data is first rate-1/2 convolutionally encoded and interleaved, then modulated by one of 64 orthogonal spreading sequences (Walsh functions). Next, a synchronized scrambling sequence unique to each cell is superimposed on top of the Walsh function to reduce interference between cells. The scrambling requires synchronization between base stations. Uplink spreading is accomplished using a combination of a rate-1/3 convolutional code with interleaving, modulation by an orthogonal Walsh function, and modulation by a nonorthogonal user/base station–specific code. The IS-95 standard includes a parity-check code for error detection, as well as power control for the reverse link to avoid the near–far problem. A three-finger RAKE receiver is also specified to provide diversity and to compensate for ISI. A form of base station diversity called soft handoff (SHO), whereby a mobile maintains a connection to both the new and old base stations during handoff and combines their signals, is also included in the standard. CDMA has some advantages over TDMA for cellular systems, including no need for frequency planning, SHO capabilities, the ability to exploit voice activity to increase capacity, and no hard limit on the number of users that can be accommodated in the system. There was much debate about the relative merits of the IS-136 and IS-95 standards throughout the early 1990s, with claims that IS-95 could achieve 20 times the capacity of AMPS whereas IS-136 could only achieve 3 times this capacity. In the end, both systems turned out to achieve approximately the same capacity increase over AMPS.

The 2G digital cellular standard in Japan, called the Personal Digital Cellular (PDC) standard, was established in 1991 and deployed in 1994. It is similar to the IS-136 standard but with 25-kHz voice channels for compatibility with the Japanese analog systems. This system operates in both the 800–900-MHz and 1400-MHz frequency bands. Each channel accommodates three users, and six users per channel can be accommodated with a higher voice compression rate at the expense of poorer voice quality.

In the late 1990s, second-generation systems evolved in two directions: they were ported to higher frequencies as more cellular bandwidth became available in Europe and the United States, and they were modified to support data services in addition to voice. Specifically, in 1994 the FCC began auctioning spectrum in the Personal Communications Service (PCS) band at 1.9 GHz (1900 MHz) for cellular systems. Operators purchasing spectrum in this band could adopt any standard. Different operators chose different standards, so GSM, IS-136, and IS-95 were all deployed at 1.9 GHz in different parts of the country, making nationwide roaming with a single phone difficult. In fact, many of the initial digital cellphones included an analog AMPS mode in case the digital system was not

available. GSM systems operating in the PCS band are sometimes referred to as PCS 1900 systems. The IS-136 and IS-95 (cdmaOne) standards translated to the PCS band go by the same names. Europe allocated additional cellular spectrum in the 1.8-GHz band. The standard for this frequency band, called GSM 1800 or DCS 1800 (for Digital Cellular System), used GSM as the core standard with some modifications to allow overlays of macrocells and microcells. Second-generation cordless phones such as Digital Enhanced Cordless Telecommunications (DECT), the Personal Access Communications System (PACS), and the Personal Handyphone System (PHS) also operated in the 1.9-GHz frequency band.

Once digital cellular became available, operators began incorporating data services in addition to voice. The 2G systems with added data capabilities are sometimes referred to as 2.5G systems. The enhancements to 2G systems made to support data services are summarized in Table D.4. GSM systems followed several different upgrade paths to provide data services. The simplest, called High Speed Circuit Switched Data (HSCSD), allows up to four consecutive timeslots to be assigned to a single user, thereby providing a maximum transmission rate of up to 57.6 Kbps. Circuit switching is quite inefficient for data, so a more complex enhancement provides for packet-switched data layered on top of the circuit-switched voice. This enhancement is referred to as General Packet Radio Service (GPRS). A maximum data rate of 171.2 Kbps is possible with GPRS when all eight timeslots of a GSM frame are allocated to a single user. The data rates of GPRS are further enhanced through variable-rate modulation and coding, referred to as Enhanced Data rates for GSM Evolution (EDGE). EDGE provides data rates up to 384 Kbps with a bit rate of 48-69.2 Kbps per timeslot. GPRS and EDGE are compatible with IS-136 as well as GSM, thus providing a convergent upgrade path for both of these systems. The IS-95 standard was modified to provide data services by assigning multiple orthogonal Walsh functions to a single user. A maximum of eight functions could be assigned, leading to a maximum data rate of 115.2 Kbps, although in practice only about 64 Kbps was achieved. This evolution was referred to as the IS-95b standard.

2G Standard	2.5G Enhancement	Technique	Max Data Rate (Kbps)	Typical Data Rate (Kbps)
GSM	HSCSD	Aggregate Timeslots	57.6	14,4-57.6
GSM/IS-136	GPRS	Aggregate Timeslots, Packet Switching	140.8	56
GSM/IS-136	EDGE	GPRS with variable modulation and coding	384	200
IS-95	IS-95b	Aggregate Walsh Functions	115	64

Table D.4: 2G enhancements to support 2.5G data capabilities

Third-Generation CDMA Systems

The fragmentation of standards and frequency bands associated with 2G systems led the International Telecommunications Union (ITU) in the late 1990s to formulate a plan for a single global frequency band and standard for third-generation (3G) digital cellular systems. The standard was named the International Mobile Telephone 2000 (IMT-2000) standard with a desired system rollout in the 2000 timeframe. In addition to voice services, IMT-2000 was to provide Mbps data rates for demanding applications such as broadband Internet access, interactive gaming, and high-quality audio and video entertainment. Agreement on a single standard did not materialize, with most countries supporting one of two competing standards: cdma2000 (backward compatible with cdmaOne and led by North America), supported by the Third Generation Partnership Project 2 (3GPP2); and wideband CDMA (W-CDMA, backward compatible with GSM and IS-136 and led by Europe and Asia), supported by the Third Generation Partnership Project 1 (3GPPl). The main characteristics of these two 3G standards are summarized in Table D.5. Both standards use CDMA with power control and RAKE receivers, but the chip rates and other specification details are different so they are not compatible standards. A third 3G standard, TD-SCDMA, was adopted in China. The key difference between TD-SCDMA and the other 3G standards is its use of TDD instead of FDD for uplink/downlink signaling.

The cdma2000 standard built on cdmaOne to provide an evolutionary path to 3G. The core of the cdma2000 standard operated in one pair of 1.25-MHz radio channels and was thus backward compatible with cdmaOne systems. The cdma2000 1X system doubled the voice capacity of cdmaOne systems and supported high-speed data with projected peak rates of about 300 Kbps and actual rates of about 144 Kbps. There were two evolutions of this core technology to provide high data rates (referred to as HDR service) above 1 Mbps, which were referred to as cdma2000 1XEV. The first phase of evolution, cdma2000 1XEV-DO (data only), enhanced the cdmaOne system using a separate 1.25-MHz dedicated high-speed data channel that supported downlink data rates up to 3 Mbps and uplink data rates up to 1.8 Mbps for an average combined rate of 2.4 Mbps. The second phase of the evolution, cdma2000 1XEV-DV (data and voice), supported up to 4.8-Mbps data rates as well as legacy 1X voice users, 1XRTT data users, and 1XEV-DO data users – and all within the same radio channel. Another enhancement to cdma2000 was to aggregate three 1.25-MHz channels into one 3.75-MHz channel.

W-CDMA, also referred to as the Universal Mobile Telecommunications System (UMTS), became the 3G successor to GSM. The W-CDMA link layer protocol was also used in the Japanese FOMA and J-Phone 3G systems with different protocols above the link layer for other aspects of the system such as routing and speech compression. W-CDMA supports peak rates of up to 2.4 Mbps, with typical rates anticipated in the 384-Kbps range. W-CDMA uses 5-MHz channels, in contrast to the 1.25-MHz channels of cdma2000. An enhancement to W-CDMA called High Speed Data Packet Access (HSDPA) provides data rates of around 9 Mbps.

3G marked the first release of the 3GPP cellular standard, beginning a cadence of regular 3GPP releases for each new cellular standard generation and the evolution of the standard between generations. In particular, UMTS standardizing W-CDMA for 3G was the first release of the 3GPP cellular standard and was labeled 3GPP Release 99. The numbering then changed, with Release 4 defining 3G core network improvements as the next release. Additional 3G standards improvements were captured in 3GPP Releases 5-7, with Release 7 significantly increasing data rates through High-Speed Packet Access (HSPA). Release 7 was the last 3GPP standards release for 3G networks, with subsequent releases focused on later generations.

System	CDMA 2000				W-CDMA		
Version	1X	1XEV-DO	1XEV-DV	3X	UMTS	FOMA	J-Phone
Channel Bandwidth (MHz)	1.25	1.25	1.25	3.75	5	5	5
Chip Rate (Mchips/sec)	1.2288	1.2288	1.2288	3.6864	3.84	3.84	3.84
Peak Data Rate (Mbps)	.144	2.4	4.8	5-8	2.4 (8-10 with HSDPA)	2.4 (8-10 with HSDPA)	2.4 (8-10 with HSDPA)
Modulation	QPSK/MPSK (Downlink); BPSK/QPSK (Uplink)				QPSK (Downlink); BPSK (Uplink)		
Coding	Convolutional (Low Rate); Turbo (High Rate)						

Table D.5: Third-generation digital cellular phone standards

Fourth-Generation OFDMA/MIMO Systems: WiMax and LTE

The introduction of the iPhone in the mid-2000s severely stressed the capabilities of 3G cellular systems, creating significant urgency to define and launch the next generation cellular standard. This led to intense activity by the ITU to define 4G cellular requirements tailored to the needs of smartphone usage, which was growing rapidly. In particular, the ITU specification for 4G included an order of magnitude higher data rates for mobile users, up to 100 Mbps for mobile users, as well as up to 1 Gbps for stationary users. Latency requirements were also reduced by over an order of magnitude relative to 3G, with a maximum of 100 ms, in order to support video streaming and gaming. A major change from 3G to 4G network designs was to change the physical layer from CDMA to OFDMA. OFDMA was far better suited to high speed data than CDMA as it allows for link layer adaptation to channel conditions and more flexible bandwidth allocation to users. Another major change from 3G to 4G was to require all-IP based network architecture tailored to high speed data applications, in contrast to the previous circuit-switched architecture tailored to voice.

Although 4G standardization to meet these 4G ITU requirements was progressing in the late 2000s for a launch around 2010, carriers could not wait for the new standard to upgrade their networks given the poor performance of smartphones on the 3G network. Hence, carriers opted to upgrade their networks to the "long term evolution (LTE)" standard (3GPP Release 8), which was designed for a relatively smooth transition from the 3G standards that had been developed by 3GPP. LTE provided better performance than 3G but did not meet minimum requirements of 4G. Simultaneously, an alternative cellular standard to LTE called mobile WiMax was developed within the IEEE standards process to meet 4G requirements [5]. The mobile WiMax standard, also known as the IEEE 802.16m standard, evolved from the IEEE 802.16 standard for fixed broadband systems and has many of the same technical features as LTE. While mobile WiMax had better performance than LTE initially, the smooth upgrade path for carriers from 3G to LTE led to the latter's widespread deployment, effectively making it the defacto cellular standard in most parts of the world.

Deployment of 4G networks based on the LTE standard (3GPP Releases 8-9) became widespread by the mid-2010s. In the late 2010s the 4G standard was enhanced with multiple antenna (MIMO) technology to include up to 4 antennas at both the base station and the mobile device, although few mobile devices had more than 2 antennas. It also introduced carrier aggregation, combining multiple channels that could be assigned to a single user. These enhancements were coined LTE Advanced and were captured in 3GPP Releases 10-14. The high data rates and low latency of LTE Advanced networks meet the requirements for many smart phone applications, including video streaming, video calls, and gaming. The key features of LTE and LTE Advanced were summarized in Table D.1.

Fifth-Generation Systems: OFDMA/MIMO, mmW, and Network Virtualization

In keeping with the cadence of a new generation of cellular standard every decade, the ITU began the process to define requirements for 5G as soon as 4G networks were launched. These 5G requirements, coined IMT-2020 for an anticipated launch of cellular systems meeting these requirements in 2020, included an order of magnitude increase in data rates and reduction in latency relative to 4G systems. Specifically, the IMT-2020 requirements for 5G specified peak data rates of up to 20 Gbp s with uplink and downlink speeds of up to 10 Gbps. In addition to the increased data rates, an important change in these requirements with respect to 4G and earlier cellular generations was that for 5G, uplink and downlink speeds were specified to be equal. That is because the 5G use cases had mobile devices generate and upload as much content as they download, whereas 4G and earlier cellphone use cases had far higher data rates for download versus upload. This posed a significant technical challenge for 5G, as the base station can accommodate a larger antenna array and has access to more power than the mobile unit, so fast downlink speeds are easier to achieve than on the uplink. In order to achieve these higher data rates on the uplink, 5G systems increased the antenna requirements from 4x4 MIMO in 4G to as large as 128x8 MIMO. They also began operating in the mmW bands, where spectrum is plentiful. However, high path loss and shadowing at those frequencies limited the areas where mmW spectrum could be used, typically within a small distance of a base station, even with the large antenna arrays specified for 5G. 5G also reduced latency from the 10-20ms of 4G to 1-4ms. This reduction required changes to the network core to enable network slicing via network virtualization. As discussed in Section 15.4.2, wireless network virtualization consists of software in the cloud or core network that creates multiple virtual networks offering different user services that run over the same physical network infrastructure. With this architecture operators can assign a dedicated slice of the end-to-end 5G network resources, i.e. a virtual network, to applications with stringent delay, speed, jitter, and other requirements. The architecture of 4G supported best-effort packet switching for all applications, so network virtualization and slicing was a significant change in the evolution from 4G to 5G. The 5G physical layer is the same as in 4G: OFDM/OFDMA. As discussed above, the MIMO antenna array size increased from 4x4 to 128x8. 5G also introduced the use of Polar codes in addition to the Turbo and LDPC codes used in 4G, primarily for the control channels due to their low decoding complexity on these short packets. The first 5G 3GPP standard with the

new radio (NR) standard was Release 15 with additional improvements in releases 16-18.

Sixth-Generation Systems and Beyond The requirements for 6G systems are currently being finalized for an expected launch of this new cellular generation in 2030. As in the 3G to 4G and 4G to 5G transitions, it is anticipated that key performance metrics for 6G will improve by approximately an order of magnitude over 5G requirements. In particular, peak data rates of up to 100 Gbps on both the uplink and downlink are expected and a latency requirement of .1-1ms is currently targeted. Once the requirements are specified, the technologies to achieve them will be incorporated into the standard. These will likely include sub Terahertz frequency bands, in the 90-300 GHz range, in addition to the 5G frequency bands currently in use. However, the limited use of mmW spectrum in 5G systems to date indicates that achieving data rates on the order of hundreds of Gigabits per second by utilizing sub-Terahertz frequency bands in practice will be very challenging. The number of base station antennas is also expected to increase dramatically in 6G versus 5G, from 128 to as many as 1028, while the number of mobile device antennas is expected to remain the same at up to 8. Preliminary work in defining the 6G standard was captured in 3GPP Release 19.

Beyond 6G, given the marketing benefits to operators and perceived if not actual benefits to users, it is likely that a new generation of the cellular standard will continue to be developed and adopted approximately every decade. In between each generation, incremental changes will be made to improve performance. Whether each new generation of cellular standard can increase performance in multiple dimensions by an order of magnitude or more, as has happened until now, will depend on whether technology advances can deliver those kinds of gains. While there are surely more wireless technology breakthroughs on the horizon, the technology has matured dramatically over the last few standards generations, particularly through the use of OFDM, massive MIMO, and network virtualization. If new technology breakthroughs arise more slowly than in the past, we might see limited performance gains as we go from, say, 6G to 7G or 7G to 8G. In addition, it is not clear that compelling current and future use cases require significantly better performance than is possible with today's cellular technology, or whether these use cases would generate enough revenue to merit the cost to operators of deploying an entirely new generation of cellular technology across their network. These are open questions that will be answered over the next few decades.

D.2 Wi-Fi Standards

Wi-Fi standards are based on the family of IEEE 802.11 standards, developed by the IEEE through its 802.11 working group. The Wi-Fi acronym is an abbreviation for wireless fidelity. A summary of the key features for each generation of Wi-Fi standard is summarized in Table D.6 .

As indicated in the table, the first 802.11 Wi-Fi standard was launched in 1997, but the first commercial Wi-Fi networks to be developed were based on the 802.11b standard adopted in 1999. The 11 Mbps data rate of this standard, operating in the 2.4 GHz frequency band with BPSK/QPSK, DSSS, and simple channel codes, was several orders of magnitude higher than the data rates of the 3G cellular launched the following year, primarily due to the much larger channel bandwidth for Wi-Fi (20 MHz) versus cellular (1.25 MHz). Moreover, the performance of 802.11b was close to the speed and reliability of the Ethernet wireed LANs at the time. Achieving the same network performance as an Ethernet connection without the need for wired connectivity led to fast and wide adoption of Wi-Fi by consumers and businesses. In particular, millions of Wi-Fi access points were sold in the first year after 802.11b products were launched. This led a broad range of consumer devices, including laptops and cellular phones, to incorporate Wi-Fi chipsets. Indeed, the success of 802.11b and subsequent Wi-Fi standards has led to one of the most ubiquitous technologies available today; Wi-Fi network access is available in almost every corner of the globe, and most consumer devices are Wi-Fi enabled to access these networks.

The 802.11a standard, operating in the 5 GHz frequency band, was finalized at the same time as 802.11b and

Standard	802.11	802.11a	802.11b	802.11g	802.11n (Wi-Fi 4)	802.11ac (Wi-Fi 5)	802.11ad	802.11ah	802.11ax (Wi-Fi 6)	802.11ay	802.11be (Wi-Fi 7)
Release Date	1997	1999	1999	2003	2009	2013	2018	2017	2021	2021	2024
Frequency Bands (GHz)	2.4	5	2.4	2.4	2.4/5	5	60	<1	2.4/5/6	60	2.4/5/6
Channel Bandwidth (MHz)	20	20	20	20	20,40	20,40,80,160	2160	1, 2, 4, 8, 16	20,40,80,160	2.16,4,32, 6.48, 8.64	Up to 320
Max. Data Rate	2 Mbps	54 Mbps	11 Mbps	54 Mbps	600 Mbps	6.9 Gbps	8.1 Gbps	347 Mbps	9.6 Gbps	303 Gbps	46.1 Gbps
Max. Range (Indoor/Outdoor)	20m/100 m	35m/120m	35m/140m	38m/140m	70m/250m	35m indoor	3.3m indoor	1Km outdoor	30m/120m	10m/100m	30m/120m
Multiple Access		CSMA/CA	CDMA/CA	CSMA/CA	CSMA/CA	CSMA/CA	CSMA/CA	CSMA/CA	OFDMA	OFDMA	OFDMA
Modulation	BPSK, QPSK	Adaptive 64 QAM	BPSK, QPSK	Adaptive 64 QAM	Adaptive 64 QAM	Adaptive 256 QAM	Adaptive 256 QAM	Adaptive 256 QAM	Adaptive 1024 QAM	Adaptive 1024 QAM	Adaptive 1024 QAM
Coding		BCC	CCK	BCC	BCC, LDPC	BCC, LDPC	BCC, LDPC	BCC, LDPC	BCC, LDPC	BCC, LDPC	BCC, LDPC
ISI Mitigation	DSSS, FHSS	OFDM	DSSS	OFDM	OFDM	OFDM	OFDM	OFDM	OFDM	OFDM	OFDM
MIMO; spatial streams;MU-MIMO	No;1;No	No;1;No	No;1;No	No;1;No	4x4;4;No	8x8;8;Yes	32x1;1;Yes	4x4;4;Yes	8x8;8;Yes	8x8;8;Yes	16x16,16,Yes

Table D.6: Timeline and Key Features of the 802.11 Wi-Fi Standards

offered significantly higher data rates (54 Mbps versus 11 Mbps) through a more sophisticated physical layer design (OFDM and adaptive MQAM modulation with a maximum constellation size of 64 QAM and binary convolutional codes (BCC)). However, the higher cost, higher energy consumption, shorter range, worse reliability, and need to avoid radar for Wi-Fi operating in parts of the 5 GHz band led to little adoption of 802.11a. A new Wi-Fi standard, 802.11g, was released a few years later, in 2003. The 802.11g standard is essentially the same design as 802.11a except that it operates in the 2.4 GHz band instead of the 5 GHz band. Many Wi-Fi access points and devices quickly adopted this new standard to take advantage of its higher data rates over 802.11b. Moreover, as the 2.4 GHz band became more crowded and as 5 GHz radio technology improved, 802.11a started to gain more users as well. Most Wi-Fi access points and chipsets released in this timeframe incorporated all three standards, 802.11a, 802.11b, and 802.11g, to take advantage of all available Wi-Fi bandwidth and be backwards compatible with older generations of Wi-Fi technology.

The next new Wi-Fi standard, 802.11n, released in 2009, incorporated MIMO for the first time, with up to 4 antennas at both the access point and end devices, enabling up to four independent spatial data streams. This standard operates in both the 2.4 GHz and 5 GHz frequency bands, so it can exploit the extended range of the 2.4 GHz band as well as the significantly more bandwidth available in the 5 GHz band. Channel bandwidth was also doubled relative to 802.11g, from 20 MHz to 40 MHz. In addition, optional LDPC channel coding was added as an alternative to the BCCs used in 802.11a/g. These improvements increased data rates from the 54 Mbps of 802.11g to up to 600 Mbps. As stated above, the 3G cellular standard at the time could only support 7 Mbps.

The next major improvement in Wi-Fi technology was to incorporate downlink multiuser MIMO (MU-MIMO), introduced in the 802.11ac standard released in 2013. MU-MIMO allows the access point to send different spatial streams to different users simultaneously rather than sequentially, which increases spectral efficiency and reduces latency. 802.11ac also increased the MIMO antenna configuration to up to 8x8 antennas with 8 spatial streams, and the maximum constellation size of adaptive MQAM to 256 QAM. Through these improvements, 802.11ac could achieve data rates up to 6.9 Gbps.

Subsequent generations of the Wi-Fi standard either improved upon existing techniques by adding more antennas for MIMO or increasing the maximum adaptive MQAM constellation size, or adopted new technologies such as OFDMA, assigning different subcarriers to different users to increase user diversity and reduce latency. The 802.11ad and 802.11ay standards operate exclusively in the 60 GHz mmW band to exploit the large amount of available spectrum in that band. The 802.11ah standard was designed to operate with a small channel bandwidth (less than 1 GHz) and lower data rates, targeting low-power low-rate devices over large distances outdoors.

Unlike cellular standards, there has been no competing standard to 802.11 in any generation of Wi-Fi. In addition, the standard does not follow a regular cadence in updating its standard; a new 802.11 standard is launched as soon as the working group finalizes it. The confusing nomenclature for new Wi-Fi standards indicated through letters following 802.11 was changed in 2021 with the 802.11ax standard coined Wi-Fi 6, with retroactive coining of 802.11n as Wi-Fi 4 and 802.11 ac as Wi-Fi 5. This renaming was done to mimic the nomenclature in cellular standards.

We now describe the key features of each generation of Wi-Fi standard. More details can be found in [6, 7].

- **802.11:** The baseline 802.11 standard, released in 1997, specifies 83.5 MHz of bandwidth in the unlicensed 2.4 GHz frequency band along with PSK modulation coupled with FHSS or DSSS. The standard specifies data rates of up to 2 Mbps, with CSMA/CA used for random access.

- **802.11b** The baseline 802.11 standard was expanded in 1999 to create the 802.11b standard, operating in the same 2.4 GHz band as 802.11 and using only DSSS. This standard specifies variable-rate modulation and coding, with BPSK or QPSK for modulation and channel coding via either Barker sequences or complementary code keying (CCK). The 802.11b standard supports up to a maximum channel rate of 11 Mbps and a maximum user data rate of about 1.6 Mbps. The maximum transmission range is approximately 35m indoors and 120 m outdoors. The network architecture in 802.11b is specified as either star or peer-to-peer, although the peer-to-peer feature was rarely used and subsequent Wi-Fi standards all adopted the star architecture. This standard was widely deployed shortly after its introduction and launched the Wi-Fi revolution resulting in the ubiquity of Wi-Fi availability today.

- **802.11a:** The 802.11a standard was finalized in 1999 as an extension to 802.11 to improve on the 802.11b data rates. The 802.11a standard specifies 300 MHz of spectrum in the 5-GHz U-NII band segmented into three 100-MHz subbands: a lower band (5.15–5.25 GHz), a middle band (5.25–5.35 GHz), and an upper band (5.725–5.825 GHz), which has to avoid radar operations that are the primary users in that band. Channels are spaced 20 MHz apart. Three maximum transmit power levels are specified: 40 mW for the lower band, 200 mW for the middle band, and 800 mW for the upper band. These restrictions imply that the lower band is mostly suitable for indoor applications only, the middle band for indoor and outdoor, and the high band for outdoor. Variable-rate modulation and binary convolutional coding (BCC) is used on each channel: the modulation varies over BPSK, QPSK, 16-QAM, and 64-QAM, and the code rate varies over 1/2, 2/3, and 3/4. This leads to a maximum data rate per channel of 54 Mbps. The 802.11a standard uses orthogonal frequency division multiplexing (OFDM) multiple access for ISI mitigation instead of FHSS or DSSS, and in that sense it diverges from the original 802.11 standard. Following the introduction of OFDM in 802.11a, it would remain part of all Wi-Fi standards going forward. For indoor systems, the 5-GHz carrier frequency increased path loss and made it more difficult to penetrate walls and other objects relative to 2.4 GHz signals, leading to significantly lower range and speed of 802.11a relative to 802.11b. For that reason, 802.11a was used far less often than 802.11b, except in environments when its higher data rates could be consistently achieved.

- **802.11g:** The 802.11g standard, finalized in 2003, combines the best of 802.11a and 802.11b by adopting the physical and multiple access designs of 802.11a but operating in the 2.4 GHz band. Thus, 802.11g provides data rates of up to 54 Mbps with similar range and reliability as 802.11b.

- **802.11n/Wi-Fi 4:** The 802.11n standard, introduced in 2009, added several new features relative to 802.11g. It doubles the possible channel bandwidth such that either 20 MHz or 40 MHz channels can be used. It operates in both the 2.4 GHz band of 802.11b/g and the 5 GHz band of 802.11a. It also uses MIMO technology to improve robustness through diversity, particularly in the 5 GHz band, and to increase data

rates through multiplexing. In particular, it specifies a maximum 4x4 antenna configuration (4 at the access point, 4 at the client) and up to 4 independent spatial data streams where each stream uses the same adaptive modulation as 802.11g, with a maximum constellation size of 64 QAM. In addition to the convolutional codes of 802.11g, 802.11n introduces the use of quasi-cyclic LDPC coding with blocklength sizes 648, 1296, and 1944 bits as well as code rates of 1/2, 2/3, 3/4, and 5/6. All of these upgrades result in a maximum possible data rate of 600 Mbps, more than an order of magnitude higher than that of 802.11g. The range of 802.11n was double that of prior Wi-Fi standards, and its reliability was also greatly improved due to the benefits of MIMO. The 802.11n standard was retroactively named Wi-Fi 4 by the Wi-Fi Alliance in 2018.

- **802.11ac/Wi-Fi 5:** The 802.11ac standard was introduced in 2013. The primary goal of this new standard was to support Gbps data rates indoors at a range of up to 35 m. To do so, it operates only in the 5 GHz band where more spectrum is available than in the 2.4 GHz band. To take advantage of that spectrum, 802.11ac utilized channel bandwidths of 20, 40, 80, and 160 MHz. The maximum constellation size of the adaptive modulation is also increased from the 64 QAM of prior generations to 256 QAM. The MIMO maximum antenna configuration was increased from 4x4 to 8x8, with up to 8 independent spatial streams. Another new feature of 802.11ac was to use Multiuser MIMO (MU-MIMO) on the downlink, where different spatial streams can be sent from the access point to different users, which exploits user diversity and reduces downlink latency. The uplink was still restricted to sequential user by different users and the same BCC and optional QC-LDPC codes as in the 802.11n standard are used. These improvements over 802.11n support a maximum data rate of 6.9 Gbps at a maximum range of 35 m indoors, similar to the range of prior Wi-Fi standards.

- **802.11ad:** The 802.11ad standard, introduced in 2018, was the first Wi-Fi standard to utilize the 60 GHz unlicensed millimeter wave band with a single channel bandwidth of 2.16 GHz. Up to 16 multiple antennas at the access point are used to mitigate the path loss issues in this band, and downlink MU-MIMO allows the access point to simultaneously transmit to up to 16 users. The standard does not support multiple antennas at the client. The adaptive modulation and coding is the same as in 802.11ac and the maximum data rate is 8.1 Gbps. Despite the 16x1 MIMO antenna configuration, the 802.11ad standard suffers from limited range and poor reliability, hence it rarely used except in niche applications requiring high speed over short distances.

- **802.11ah:** The 802.11ah standard, introduced in 2019, was designed to support low-rate low-power applications over large outdoor areas. It operates in frequency bands below 1 GHz toimprove range, with channel bandwidths ranging from 1 MHz to 16 MHz. The standard uses the same adaptive modulation and coding as 802.11ac, the 4x4, 4 stream MIMO configuration of 802.11n, and the downlink MU-MIMO capabilities of 802.11ac. It provides a maximum data rate of 347 Mbps, far below the Gbps rates of 802.11n and ac, and a range of up to 1 Km, about four times higher than that of 802.11n.

- **802.11ax/Wi-Fi 6:** The next major leap forward for Wi-Fi was captured by the 802.11ax standard, finalized in 2021. Since the Wi-Fi Alliance in 2018 changed the naming of Wi-Fi standards to generations rather than letters affixed to the 802.11 prefix, this standard was called Wi-Fi 6 throughout its development, indicating it was the next generation after Wi-Fi 5 (802.11ac). This standard uses the 2.4 GHz and 5 GHz bands of prior Wi-Fi standards along with a new frequency band in the 6 GHz range, specifically 5.925–7.125 GHz, with the same channel bandwidths ranging from 20 MHz to 160 MHz that is used in 802.11ac. The maximum constellation size of the adaptive modulation is also increased from the 256 QAM of 802.11ac to 1024 QAM. The standard includes MIMO configurations up to 8x8 with 8 independent spatial streams and downlink MU-MIMO. It uses the same BCC and QC-LDPC codes of 802.11ac, with QC-LDPC coding mandatory for 1024 QAM, channel bandwidths greater than 20 MHz, or more than 4 spatial streams. Another new feature of 802.11ax is Orthogonal Frequency Division Multiple Access (OFDMA), where different subcarriers for

each channel can be assigned to different users on the uplink, thereby allowing simultaneous Wi-Fi uplink transmissions for the first time. The maximum data rate of 802.11ax is 9.6 Gbps with a range specified as 30 m indoors and 100 m outdoors.

- **802.11ay:** The 802.11ay standard, finalized in 2021, is an incremental upgrade of the 60 GHz 802.11ah standard. It operates in the same frequency band but allows for channel bandwidths of 2.16, 2.16,4,32, 6.48, and 8.64 GHz. It increases the maximum constellation size of the adaptive modulation to 1024 MQAM and the MIMO configuration to 8x8 with 8 independent streams and downlink MU-MIMO. OFDMA is also added. The maximum data rate of 802.11ay is 303 Gbps with a specified range of 10 m indoors and 100 m outdoors. Like 802.11ah, the challenges of operating in the 60 GHz band has limited adoption of this standard.

- **802.11be/Wi-Fi 7:** The 802.11be standard, finalized in 2024, provided incremental improvements to several features of 802.11ax/ In particular, the MIMO configuration was increased from 8x8 with 8 independent data streams to 16x16 with 16 independent data streams. Channel bandwidth was increased to include 320 MHz channels. The maximum constellation size of the adaptive MQAM was increased to 4096. All other features remain the same as in 802.11ax. The maximum data rate of 802.11be is 46.1 Gbps.

The next Wi-Fi standard, coined 802.11bn or Wi-Fi 8, is expected to be finalized in 2028. The new standard is likely to improve upon Wi-Fi 7 by increasing the number of antennas in the MIMO configuration and the number of independent spatial streams, as well as the maximum MQAM constellation. The standard may operate in the mmW band in addition to 2.4, 5, and 6 GHz bands of Wi-Fi 7, which would incorporate 802.11ah and 802.11ay into this new standard. In addition, Wi-Fi 8 is expected to introduce several new features in addition to incremental improvements to Wi-Fi 7 in order to increase reliability, lower latency and jitter, lower energy consumption, and enable more seamless handoff, load balancing, and interference mitigation between access points (APs) through multi-AP coordination. These improvements are expected to not only impact Wi-Fi 8, but also provide the foundation that will be built upon in future Wi-Fi standards.

D.3 Short-Range Wireless Network Standards

The primary standards for short-range wireless networks are Bluetooth and Zigbee, with Z-Wave a proprietary protocol that is also widely used. A summary of the main features for these three systems is shown in Table D.7 below. Bluetooth is the earliest and most ubiquitous of these systems; it offers higher data rates than the other two systems but at a higher cost as well as a shorter range. It includes a low-energy mode which has a similar energy consumption as Zigbee and Z-Wave, while also allowing for a higher power mode to increase performance and range. Bluetooth operates in the 2.4 GHz band, as does Zigbee. Z-Wave operates in the 900 MHz frequency band, offering a higher range than either Bluetooth or Zigbee, but at a lower data rate. Zigbee has a range and data rate in between that of Bluetooth and Z-Wave. One of the advantages of Zigbee and Z-Wave over Bluetooth is an ad-hoc networking capability that is designed to connect all devices within a smart home. By contrast, Bluetooth was primarily designed to replace cables in the point-to-point connection between two electronic devices; although it has an ad-hoc networking feature, it scales poorly and is rarely used. Both Bluetooth and Zigbee fall within the IEEE 802.15 family of standards for Wireless Personal Area Networks (WPANs); the IEEE adopted Zigbee as its 802.15.1 standard and Bluetooth as its 802.15.4 standard. Z-Wave is a proprietary standard, so systems based on this protocol require ownership or a license to use it, in contrast to the 802.15 standards which are free to use. We now describe each of these standards in more detail. A comparison of the main features in and tradeoffs between these systems can be found in [8, 9].

Standard	Bluetooth (802.15.1)	Zigbee (802.15.4)	Z-Wave
Frequency Band (GHz)	2.4	2.4	.9
Total Bandwidth (MHz)	83.5	83.5	26
Channel Bandwidth (MHz)	1	2	.3-.4
Max. Data Rate (Mbps)	1	.25	.1
Typical Range	10 m (100 m)	30 m	100 m
Interference/ISI Mitigation	FHSS	DSSS	None
Multiple Access	CSMA/CA	CSMA/CA	CSMA/CA
Modulation	GFSK	BPSK, OQPSK	FSK, GFSK, OQPSK
Networking	Subnet cluster (piconet: 8 nodes)	Star, Mesh, Tree (Single Hop)	Ad-Hoc (Multihop)

Table D.7: Key Features of Short-Range Wireless Network Standards and Protocols

The Bluetooth[1] radio and associated protocol was created by Ericsson in 1994 and later incorporated into the IEEE 802.15.4 standard. It enables the use of a short-range radio to replace connecting cables between electronic devices such as cell phones, tablets, headsets, earpieces, audio equipment, cameras, watches, and smart meters. Bluetooth has also been integrated into larger devices such as cars and medical equipment. Bluetooth is mainly for short-range communications – for example, from a laptop to a nearby printer or from a cell phone to a wireless headset. Its normal range of operation is 10 m (at 2.5-mW transmit power), and this range can be increased to 100 m by increasing the transmit power to 100 mW. The system operates in the unlicensed 2.4 GHz frequency band, so it can be used worldwide within this band. Other users of the 2.4 GHz band, including Wi-Fi, interfere with Bluetooth; frequency hopping is used to mitigate this interference. The Bluetooth standard and device certification is managed by the Bluetooth Special Interest Group. Bluetooth is integrated into a wide range of electronic devices with several billion Bluetooth-enabled devices shipped annually.

Bluetooth uses FSK modulation with Gaussian pulse shaping (Gaussian FSK or GFSK) for its data and voice channels. It supports one data channel at 723.2 Kbps. In this mode, there is a reverse channel with a data rate of 57.6 Kbps. The specification also allows up to three additional low-latency voice channels each at a rate of 64 Kbps, which are primarily used for cellphone connections with wireless headsets. These different modes result in an aggregate bit rate of approximately 1 Mbps. An enhanced data rate mode provides up to 3 Mbps, while a low-energy mode significantly reduces power consumption at the expense of range or data rate.

Bluetooth multiple access is based on frequency hopping spread spectrum (FHSS) with a carrier spacing and user bandwidth of 1 MHz. Typically, up to eighty different frequencies are used for a total bandwidth of 80 MHz. Users with different hop sequences simultaneously share the same 80 MHz bandwidth with interference between them whenever two different hop sequences land in the same 1 MHz channel simultaneously. Bluetooth radios can form small ad hoc networks of up to eight devices sharing the same logical channel (same hop sequence), which is called a *piconet*. with one node acting as a master and the rest as slaves. Time division is used for channel access,

[1]The Bluetooth standard is named after Harald I Bluetooth, the king of Denmark between 940 and 985 A.D. who united Denmark and Norway. Bluetooth proposes to unite devices via radio connections, hence the inspiration for its name.

with the master node coordinating the frequency-hopping sequence and synchronization with the slave nodes. Extended networks, or *scatternets*, can be formed when one node is part of multiple piconets. However, forming large networks through this approach is difficult owing to the synchronization requirements of FHSS. Collisions will occur when devices in different piconets that are on different logical channels happen to use the same hop frequency at the same time. As the number of piconets in an area increases, the number of collisions increases and performance degrades. The ad-hoc networking feature of Bluetooth via piconets and scatternets performs poorly and is rarely used. More details about Bluetooth can be found in Chapter 10 of [10, 11].

The ZigBee[2] radio specification is designed for lower cost and power consumption than Bluetooth. It supports data rates of up to 250 Kbps, far less than that of Bluetooth, using PSK modulation and DSSS at a range of up to 30 m. It follows the link and MAC protocols of the IEEE 802.15.4 standard as well as higher-layer protocols for ad hoc networking (mesh, star, or tree topologies), power management, and security. ZigBee operates in the same 2.4 GHz ISM band as Bluetooth. Its lower data rate relative to Bluetooth allows for lower power consumption and a larger transmission range. Zigbee also operates in a "green" mode whereby the radio is powered through energy harvesting of its environment, reducing or in some cases completely eliminating the need for battery power. Its device certification is managed by the Zigbee Alliance. ZigBee generally targets applications requiring relatively low data rates, low duty cycles, and large networks, such as smart tags, meters, lights, and thermostats, as well as those used for sensing and automation. The goal is to allow such devices to operate for months or years without recharging. More details about Zigbee can be found in [12, 13].

Z-Wave radios are designed primarily for smart home applications with operation in the 900 MHz ISM band. The specific frequencies in this band depend on the country of operation. Since power falls off more slowly in this band than at the higher 2.4 GHz spectrum, Z-Wave radios have a higher range than standard Zigbee or Bluetooth radios, on the order of 50-100 m. In addition, there is no interference between Z-Wave radios and those operating in the crowded 2.4 GHz band, including Wi-Fi as well as Bluetooth and Zigbee. On the downside, Z-Wave has significantly lower data rates than either Bluetooth or Zigbee, ranging from 10 to 100 Kbps. This range of data rates is supported by 3 different modulation schemes: FSK, GFSK, and OQPSK. A more detailed summary of Z-Wave technology and protocols can be found in [14]. Z-Wave was originally developed as a proprietary protocol by Zensys, and its ownership eventually transitioned to Silicon Labs. While it was a strictly proprietary technology, it later adopted a hybrid model where some aspects of the specification are open to modification by members of the Z-Wave Alliance, which manages the standard and device certification. Due to its proprietary elements, Z-Wave is not a candidate to be adopted into the 802.15 family of standards.

[2]ZigBee takes its name from the dance that honey bees use to communicate information about newly found food sources to other members of the colony.

Bibliography

[1] M. Sauter, *From GSM to LTE-Advanced Pro and 5G: An Introduction to Mobile Networks and Mobile Broadband*, 3rd ed., Wiley, 2017.

[2] W. Jiang and F.-L. Luo, *6G Key Technologies: A Comprehensive Guide*, Wiley-IEEE Press, 2022.

[3] M. Z. Chowdhury, M. Shahjalal, S. Ahmed and Y. M. Jang, " 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions", *IEEE Open Journal of the Communications Society*, 2020.

[4] M. Bakni, Wikipedia Cellular Network Standards and Generation Timeline. CC BY-SA 4.0, via Wikimedia Commons.

[5] S. Ahmadi, *Mobile WiMAX, A Systems Approach to Understanding IEEE 802.16m Radio Access Technology*, Wiley, 2011.

[6] IEEE Standards Organization, "The Evolution of Wi-Fi Technology and Standards", May 2023 (https://standards.ieee.org/beyond-standards/the-evolution-of-wi-fi-technology-and-standards/).

[7] G. Cheng, Z. Qian, Y. Jiang, Z. Yang, and T. Rajamanickam, *Wi-Fi 7: Principles, Technology, and Applications*. Springer Nature Singapore, 2025.

[8] S. J. Danbatta and A. Varol, "Comparison of Zigbee, Z-Wave, Wi-Fi, and Bluetooth Wireless Technologies Used in Home Automation," *Proc. 2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, Barcelos, Portugal, 2019, pp. 1-5.

[9] A. B. Varol, "Compilation of Data Link Protocols: Bluetooth Low Energy (BLE), ZigBee and Z-Wave," *Proc. 2019 4th International Conference on Computer Science and Engineering (UBMK)*, Samsun, Turkey, 2019, pp. 85-90.

[10] M. Hassan, *Wireless and Mobile Networking*, CRC Press, 2022.

[11] K.S. Mohamed, *Bluetooth 5.0 Modem Design for IoT Devices*, Springer, 2022.

[12] S. Farahani, *ZigBee Wireless Networks and Transceivers*. Newnes, 2011.

[13] C. Wang, T. Jiang, and Q Zhang, ZigBee Network Protocols and Applications, CRC Press, 2014.

[14] C. Paetz, *Z-Wave Essentials*. eBook Partnership; 2017.