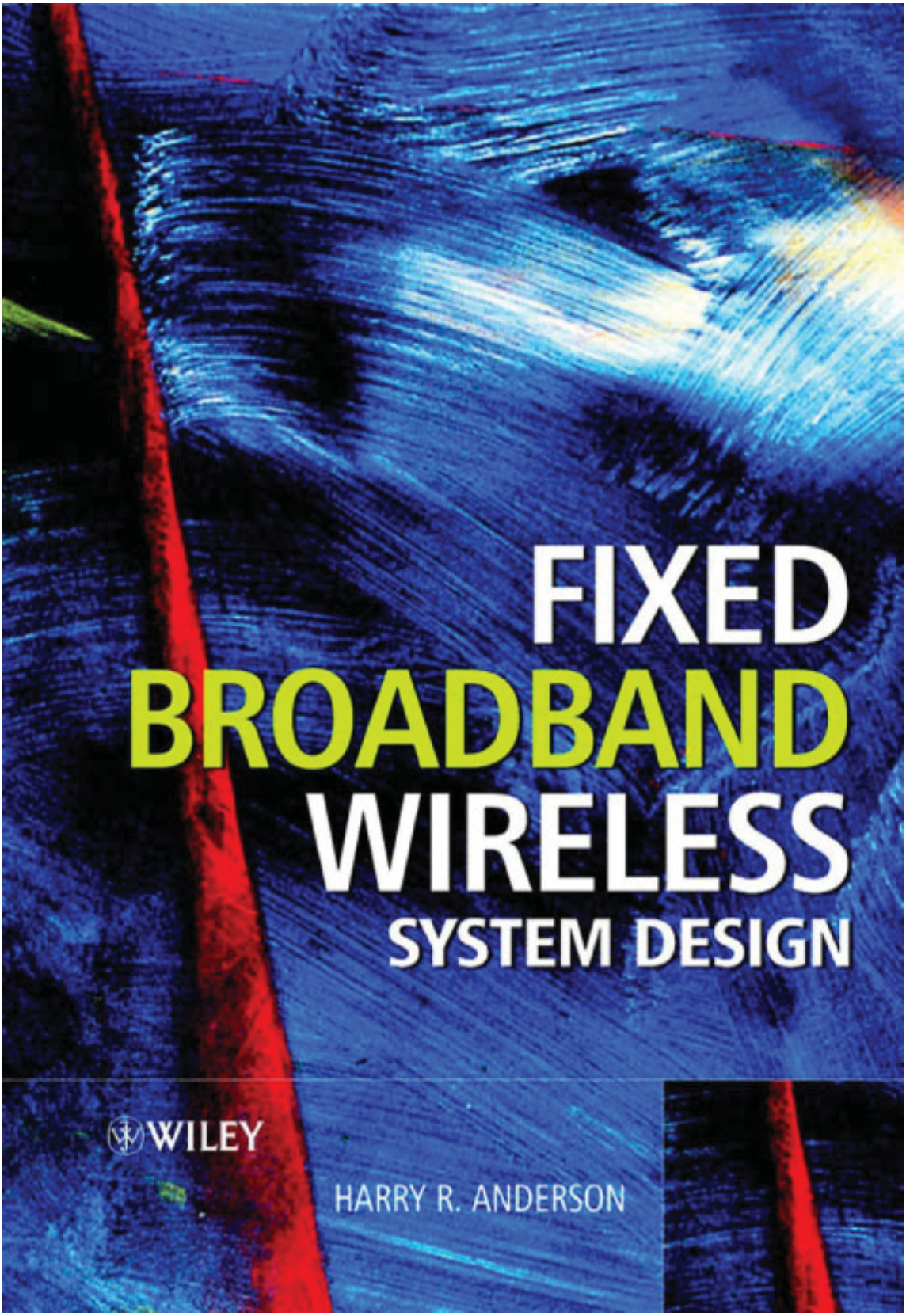
FIXED
BROADBAND
WIRELESS
SYSTEM DESIGN
WILEY
HARRY R. ANDERSON

FIXED BROADBAND WIRELESS

SYSTEM DESIGN

FIXED BROADBAND WIRELESS
SYSTEM DESIGN

HARRY R. ANDERSON, Ph.D., P.E.

CONSULTING ENGINEER
USA

WILEY

Copyright © 2003

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770571.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Anderson, Harry R.
Fixed broadband wireless system design / Harry R. Anderson.
p. cm.
Includes bibliographical references and index.
ISBN 0-470-84438-8 (alk. paper)
1. Wireless communication systems – Design and construction. 2. Cellular telephone systems – Design and construction. 3. Broadband communication systems. I. Title.

TK5103.4 .A53 2003
621.3845′6 – dc21

2002033360

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-84438-8

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Biddles Ltd, Guildford, Surrey
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Preface **xvii**

1 Fixed Broadband Wireless Systems **1**

1.1 Introduction 1
1.2 Evolution of Wireless Systems 2
1.3 Models for Wireless System Design 4
1.4 Demand for Communication Services 7
1.5 Licensed Frequency Bands 8
1.6 License-Exempt Bands 10
1.7 Technical Standards 12
1.7.1 IEEE 802.11 Standards 13
1.7.2 IEEE 802.16 Standards 14
1.7.3 ETSI BRAN Standards 14
1.8 Fixed, Portable, and Mobile Terminals 15
1.9 Types of Fixed Wireless Networks 17
1.9.1 Point-to-Point (PTP) Networks 17
1.9.2 Consecutive Point and Mesh Networks 17
1.9.3 Point-to-Multipoint (PMP) Networks 18
1.9.4 NLOS Point-to-Multipoint Networks 19
1.10 Organization of this Book 20
1.11 Future Directions in Fixed Broadband Wireless 22
1.12 Conclusions 23
1.13 References 23

2 Electromagnetic Wave Propagation **25**

2.1 Introduction 25
2.2 Maxwell's Equations and Wave Equations 25
2.3 Plane and Spherical Waves 27
2.3.1 Impedance of Free Space and Other Transmission Media 28
2.3.2 Power in a Plane Wave 29
2.3.3 Spherical Waves 29
2.4 Linear, Circular, Elliptical, and Orthogonal Polarizations 30
2.5 Free-Space Propagation 31
2.5.1 Path Attenuation between Two Antennas 31
2.5.2 Field Strength at a Distance 32

2.6 Reflection 33
2.6.1 Specular Reflection 33
2.6.2 Physical Optics 35
2.6.3 Reflections from Rough Surfaces 37
2.7 Diffraction 40
2.7.1 Wedge Diffraction 40
2.7.2 Knife-Edge Diffraction 45
2.8 Fresnel Zones and Path Clearance 51
2.9 Material Transmission 53
2.9.1 Transmission into Structures 54
2.9.2 Transmission through Foliage 54
2.10 Atmospheric Refraction 56
2.10.1 Statistics of Varying Refractivity Gradients 59
2.10.2 Sub-Refraction 61
2.10.3 Super-Refraction and Ducting 61
2.11 Atmospheric Absorption 62
2.12 Rain Attenuation and Depolarization 62
2.13 Free-Space Optics (FSO) Propagation 65
2.13.1 Beam Divergence 66
2.13.2 Fog, Snow, and Rain Attenuation 67
2.13.3 Atmospheric Scintillation 67
2.14 Conclusions 68
2.15 References 68

3 Propagation and Channel Models 71

3.1 Introduction 71
3.1.1 Model Classifications 72
3.1.2 Fading Models 73
3.2 Theoretical, Empirical, and Physical Models 73
3.2.1 Theoretical Channel Models 74
3.2.1.1 Theoretical, Non-Time-Dispersive 75
3.2.1.2 Theoretical, Time-Dispersive 75
3.2.2 Empirical Channel Models 75
3.2.2.1 Empirical, Non-Time-Dispersive 77
3.2.2.2 Empirical, Time-Dispersive 77
3.2.3 Physical Channel Models 78
3.2.3.1 Physical, Non-Time-Dispersive, Not Site-Specific 78
3.2.3.2 Physical, Non-Time-Dispersive, Site-Specific 78
3.2.3.3 Physical, Time-Dispersive, Site-Specific 79
3.3 Generic Wideband Channel Model 79
3.3.1 Wideband Channel Response 83
3.3.1.1 Time–Variant and Static Channels 85

3.3.1.2 Tapped Delay Line Model 88
3.3.1.3 Frequency Domain Representations 89
3.4 Empirical Models 89
3.4.1 IEEE 802.16 (SUI) Models 90
3.4.2 COST-231 Hata Model 93
3.4.3 MMDS Band Empirical Path Loss 94
3.4.4 3D Path Loss Surface Models 96
3.5 Physical Models 97
3.5.1 Free Space + RMD 98
3.5.1.1 Line-of-Sight Assessment 98
3.5.1.2 LOS Path Analysis 99
3.5.1.3 NLOS Path Analysis 102
3.5.2 Multiple Obstacle Analysis 102
3.5.2.1 Epstein–Peterson Method 105
3.5.2.2 Deygout Method 106
3.5.3 Longley–Rice Model 107
3.5.4 TIREM Model 107
3.5.5 Anderson 2D Model 107
3.5.6 NLOS Dominant Ray Path Loss Model 108
3.5.6.1 Building Clutter Loss 110
3.5.7 Ray-Tracing 115
3.5.8 Simplified Indoor Model 120
3.6 Conclusions 122
3.7 References 123

4 Fading Models 127

4.1 Introduction 127
4.1.1 Link Performance with Fading 128
4.2 Atmospheric Fading Models 129
4.2.1 Microwave Multipath Fading Mechanisms 130
4.2.2 Vigants–Barnett Model 132
4.2.3 ITU-R P.530-8 Model 134
4.2.4 Dispersive (Frequency-Selective) Fading 137
4.2.4.1 Coherence Bandwidth 138
4.2.4.2 Dispersive Fade Margin 140
4.3 Rain Fading Models 143
4.3.1 Crane Rain Fade Model 144
4.3.2 ITU-R P.530-8 Model 146
4.3.3 Short-Range Rain Fading 147
4.3.4 Other Precipitation Losses 149
4.3.5 Cross-Polarization Discrimination Fading Model 149
4.4 Correlated Rain Fading Model 151
4.5 Free Space Optics Fog Fading Models 152
4.6 Fading Models for NLOS Links 153

4.6.1 NLOS Multipath Fading Models 154
4.6.1.1 Rayleigh Distribution 154
4.6.1.2 Rician Distribution 157
4.6.1.3 Nakagami Distribution 158
4.6.2 NLOS Shadow Fading Models 160
4.6.3 Composite Fading–Shadowing Distributions 161
4.7 Conclusion 162
4.8 References 163

5 Propagation Environment Models 165

5.1 Introduction 165
5.2 Topography 166
5.2.1 Topographic Maps 166
5.2.2 Terrain DEMs 167
5.2.3 DEM Data from Satellite and Aerial Imagery 169
5.3 Buildings and Other Structures 171
5.3.1 Vector Building Databases 172
5.3.2 Canopy Building Databases 173
5.3.3 System Analysis Errors from Using Canopy Databases 175
5.4 Morphology (Land Use/Land Cover or Clutter) 178
5.5 Atmospheric and Meteorology Factors 179
5.5.1 Atmospheric Refractivity 180
5.5.2 Rain Rates 180
5.5.3 Fog Data 181
5.6 Mobile Elements of the Propagation Environment 181
5.7 Mapping Fundamentals 182
5.7.1 Spheroids, Ellipsoids, and Geoids 183
5.7.2 Geodetic Systems, Datums, and Datum Transformations 183
5.7.3 Map Projections 186
5.7.4 Coordinate Systems 187
5.8 Conclusions 187
5.9 References 188

6 Fixed Wireless Antenna Systems 189

6.1 Introduction 189
6.2 Antenna System Fundamentals 190
6.2.1 Radiation from an Elemental Dipole Antenna 192
6.2.2 Directivity and Gain 194
6.2.3 Antenna Radiation Patterns 195
6.2.4 Polarization 197
6.2.5 Antenna Efficiency and Bandwidth 199
6.2.6 Electrical Beamtilt, Mechanical Beamtilt, and Null Fill 201
6.2.7 Reciprocity 202

6.3 Fixed Narrow Beam Antennas 202
6.3.1 Horn Antennas 203
6.3.2 Parabolic and Other Reflector Antennas 205
6.4 Fixed Broad Beam Antennas 208
6.4.1 Horn Antennas for Hub Sectors Above 10 GHz 209
6.4.2 Hub Sector Antennas for MMDS and U-NII Bands 209
6.4.2.1 Linear Antenna Arrays 210
6.4.2.2 Planar Antenna Arrays 212
6.5 Diversity Antenna Systems 214
6.5.1 Empirical Microwave Link Diversity Improvement 217
6.6 Adaptive Antennas 217
6.6.1 Optimum Combining 219
6.7 MIMO Antenna Systems 223
6.8 Waveguides and Transmission Lines 226
6.8.1 Waveguides 227
6.8.2 Transmission Lines 228
6.9 Radomes 229
6.10 Engineered and *Ad Hoc* Antenna Installations 231
6.11 Conclusions 232
6.12 References 233

7 Modulation, Equalizers, and Coding 235

7.1 Introduction 235
7.2 Digital Modulation – Amplitude, Frequency, and Phase 236
7.3 Fixed Broadband Wireless Modulation Methods 237
7.3.1 BPSK, QPSK, $\pi/4$-DQPSK 238
7.3.2 16QAM, 64QAM, and 256QAM 239
7.3.3 Orthogonal Frequency Division Multiplexing (OFDM) 241
7.3.3.1 OFDM Peak-to-Average Power Ratio 244
7.4 Error Performance with Noise and Interference 244
7.4.1 Error Performance with Gaussian Noise Only 245
7.4.2 Error Performance with Noise and Constant Amplitude Interference 248
7.4.2.1 16QAM with Noise and Interference 250
7.4.2.2 16QAM with 16QAM Interference 253
7.4.2.3 Coherent QPSK with Noise and Interference 253
7.4.2.4 Differential QPSK with Noise and Interference 256
7.4.3 Error Performance with Flat-Fading Signal and Interference 256
7.4.3.1 Noise Approximation of Interference 257
7.4.4 Error Performance with Frequency Selective Signal Fading 257
7.5 Equalizers 259
7.5.1.1 Time Domain Symbol Equalizers 259
7.5.1.2 Frequency Domain Equalizers (FDE) 261

7.6 Coding Techniques and Overhead 262
7.6.1 Block Codes 263
7.6.1.1 Cyclic Codes 264
7.6.2 Concatenated Codes 265
7.6.3 Interleaving 265
7.6.4 Convolutional Codes 266
7.6.5 Trellis-Coded Modulation (TCM) 267
7.6.6 Coding Gain 268
7.6.7 Space-Time Codes 269
7.7 Conclusion 272
7.8 References 273

8 Multiple-Access Techniques 275

8.1 Introduction 275
8.1.1 Intersystem Multiple Access 276
8.1.2 Intrasystem Multiple Access 277
8.1.3 Duplexing 277
8.2 Frequency Division Multiple Access (FDMA) 278
8.2.1 FDMA Interference Calculations 280
8.2.1.1 Noise Power 280
8.2.1.2 Cochannel and Adjacent Channel Interference 282
8.2.1.3 Multiple Interferers in LOS Networks 284
8.2.2 Spectrum Utilization 286
8.3 Time Division Multiple Access (TDMA) 286
8.3.1 TDMA Intercell Interference 288
8.4 Code Division Multiple Access (CDMA) 290
8.4.1 Frequency-Hopping Spread Spectrum (FHSS) 290
8.4.2 Direct Sequence (DS) Spread Spectrum 291
8.4.3 Downlink Interference Calculations 293
8.4.3.1 Downlink Pilot Channel E_c/I_0 294
8.4.3.2 Downlink Traffic Channel E_b/N_0 294
8.4.4 Uplink Interference Calculations 295
8.4.4.1 Rake Receiver 297
8.4.5 Joint (Multiuser) Detection 298
8.4.6 CDMA Broadband Standards 299
8.5 Space Division Multiple Access (SDMA) 302
8.6 Carrier Sense Multiple Access (CSMA) 304
8.7 Multiple Access with OFDM 305
8.7.1 Multicarrier CDMA (MC-CDMA) 306
8.7.2 Orthogonal Frequency Division Multiple Access (OFDMA) 307
8.7.3 OFDM with TDMA 307
8.7.4 OFDM with CSMA/CA (IEEE 802.11a) 308

8.7.5 OFDM with SDMA 308
8.7.6 OFDM Multiple-Access Standards 308
8.8 Duplexing Methods 309
8.8.1 Frequency Division Duplexing (FDD) 310
8.8.2 Time Division Duplexing (TDD) 311
8.8.2.1 TDD Interference Calculations 312
8.9 Capacity 313
8.9.1 Shannon Theoretical Channel Capacity 314
8.9.2 Capacity in Interference-Limited, Multiuser Systems 315
8.9.3 User Capacity 318
8.9.4 Commercial Capacity 318
8.10 Conclusion 319
8.11 References 319

9 Traffic and Application Mix Models 321

9.1 Introduction 321
9.2 Traffic Geographic Distribution Models 323
9.2.1 Residential Demographic Data 323
9.2.2 Business Demographic Data 326
9.2.3 Land Use Data 329
9.2.4 Building Data 330
9.2.5 Aerial Photographs 331
9.3 Service and Application Types 332
9.4 Circuit-Switched Traffic Models 333
9.4.1 Circuit-Switched Quality of Service (QoS) 334
9.4.1.1 Erlang B Blocking Probability 334
9.4.1.2 Erlang C Blocking Probability 335
9.5 Packet-Switched Traffic Models 335
9.5.1 Self-Similar Data Characteristics 337
9.5.2 Packet Probability Distributions 338
9.5.2.1 Packet Size Distribution 338
9.5.2.2 Packets and ADU's 339
9.5.2.3 Packet Interarrival Time Distribution 339
9.5.2.4 Distribution of the Number of Packets and the Packet Sessions 340
9.5.2.5 Packet Session Interval Distribution 340
9.5.2.6 Packet Session Arrival Distribution 341
9.5.3 ETSI Web-Browsing Packet Transmission Model 342
9.5.4 Random Packet Cluster Transmission Model 342
9.6 Multisource Traffic Density Models 342
9.6.1 Aggregate Data Rate Statistics 344
9.6.2 Aggregate Data Rate Statistics with Packet Queuing (Delay) 346
9.6.2.1 Internet Latency 349

9.6.3 Throughput 349
9.7 Application Mix 351
9.8 Broadcast and On-Demand Video Applications 353
9.9 Conclusions 354
9.10 References 355

10 Single and Multilink System Design 357

10.1 Introduction 357
10.2 Long-Range LOS Links over Mixed Paths 358
10.2.1 Path Profile Clearance Analysis 359
10.2.1.1 Path Clearance Validation 361
10.2.2 Reflection Point Analysis 361
10.2.3 Link Budget 363
10.2.4 Fade Margin 368
10.2.5 Link Availability (Reliability) 369
10.2.6 Multipath Fade Outage 369
10.2.7 Diversity Improvement in Flat Fading Links 371
10.2.7.1 Space Diversity 371
10.2.7.2 Polarization Diversity 373
10.2.8 Dispersive (Frequency-Selective) Fade Margin 374
10.2.9 Diversity Improvement for Dispersive (Frequency-Selective) Channels 375
10.2.9.1 Frequency Diversity 376
10.2.9.2 Angle Diversity 376
10.2.10 Rain Fade Outage 376
10.2.10.1 Link Availability with Crane Rain Fade Model 377
10.2.10.2 Link Availability with the ITU-R Rain Fade Model 377
10.2.11 Composite Link Availability 378
10.2.12 Equipment Failures 379
10.3 Short-Range LOS Links in Urban Environments 380
10.3.1 Building Path Profiles 380
10.3.2 Short-Range Fading 381
10.3.3 Short-Range Urban Rain Fading 383
10.3.4 Interference Diffraction Paths over Building Edges 384
10.3.5 Urban Link Availability 385
10.3.6 Free Space Optic (FSO) Link Design 386
10.3.7 'Riser' and FSO Backup Links 387
10.4 NLOS Links in Urban and Residential Environments 387
10.4.1 Basic NLOS Path Loss 389
10.4.2 Antenna Gain in Scattering Environments 391
10.4.3 Location Variability 392
10.4.4 Time Variability (Narrowband Fading) 393

10.4.5 Time Dispersion and Arrival Angles 393
10.4.6 Channel Spatial Correlation 395
10.5 Link Adaptation 396
10.6 Multihop (Tandem) Link Systems 397
10.6.1 Passive Repeaters 398
10.7 Consecutive Point Networks 400
10.8 Mesh Networks 401
10.8.1 NLOS Mesh Networks 403
10.9 Conclusions 404
10.10 References 405

11 Point-to-Multipoint (PMP) Network Design 407

11.1 Introduction 407
11.2 LOS Network Design 409
11.2.1 Hub Site Selection 410
11.2.1.1 Visibility/Shadowing Analysis 410
11.2.1.2 Algorithms for Efficient Multiple Hub Site Selections 413
11.2.1.3 Hub Traffic/Revenue Potential Assessment 415
11.2.2 Hub Sector Configuration 416
11.2.3 CPE Best Server Hub Sector Assignments 420
11.2.4 Signal Distribution from a Rooftop 423
11.3 LOS Network Performance Analysis 423
11.3.1 Interference Analysis 424
11.3.1.1 Reduced Cross-Polarization Discrimination During Rain Fades 425
11.3.1.2 Correlated Rain Fades 425
11.3.1.3 Uplink Interference Calculations 426
11.3.1.4 Impact of Automatic Power Control (APC) 427
11.3.1.5 Coupled Links 427
11.3.2 Estimating Hub Sector Capacity Requirements 428
11.3.3 LOS Network Performance Statistics 431
11.4 NLOS Network Design 432
11.4.1 NLOS Hub Site Selection 432
11.4.1.1 Coverage/Service Area Calculations 432
11.4.1.2 Automatic Algorithms for Hub Site Selections 434
11.4.2 CPE Locations 435
11.5 NLOS Network Performance Analysis 435
11.5.1 Downlink Signals for Basic NLOS Interference Analysis 436
11.5.1.1 Downlink Interference Analysis 436
11.5.1.2 Uplink Interference Analysis 438
11.5.2 Dynamic Monte Carlo Interference Simulation 439
11.5.3 Estimating Hub Sector Capacity Requirements 442

11.5.4 NLOS Network Performance Statistics 443
11.5.5 W-CDMA Interference and Capacity 444
11.6 Network Design Revisions 444
11.6.1 PMP Network Coverage Deficiencies 445
11.6.2 High Frame Error Rates 445
11.6.3 High Packet Delay Times 445
11.7 Conclusion 446
11.8 References 447

12 Channel Assignment Strategies 449

12.1 Introduction 449
12.2 Frequency, Time Slot, and Code Planning 451
12.3 Fixed Assignments for Point-to-Point LOS Networks 451
12.3.1 Multiple Interferers on a Channel 454
12.3.2 Impact of Automatic Power Control (APC) 455
12.4 Fixed Assignments for LOS PMP Networks 455
12.4.1 LOS Networks 455
12.4.2 Conventional Cluster Frequency Planning 459
12.4.3 Impact of Adaptive Antennas in Fixed LOS Networks 460
12.4.4 Demand-Based Fixed LOS Assignments 461
12.4.5 Number of CPEs Supported in Fixed LOS Networks 464
12.5 Fixed Assignments for NLOS PMP Networks 464
12.5.1 Target S/I Ratio 465
12.5.2 Frequency Reuse Distance 466
12.5.3 Cell Layout and Channel Assignment Patterns 467
12.6 Optimizing Channel Assignments in NLOS Networks 469
12.6.1 Steepest Descent Method 470
12.6.2 Simulated Annealing Method (SA) 471
12.6.3 Genetic or Evolutionary Algorithm Method 471
12.6.4 Channel Assignments in W-CDMA Systems 472
12.7 NLOS Network Capacity 472
12.8 Dynamic Frequency Channel Assignments 473
12.8.1 Centralized DCA 473
12.8.2 Decentralized DCA 474
12.8.3 Channel Segregation 475
12.8.4 Dynamic Packet Assignment 476
12.8.5 DCA for UTRA-TDD Networks 477
12.9 Other Capacity Enhancement Techniques 478
12.9.1 Adaptive Antennas 478
12.9.2 Joint Detection 478
12.9.3 Link Adaptation 479
12.10 Spectrum Vectors, Occupancy, and Utilization 479
12.10.1 Spectrum Vectors 480
12.10.2 Spectrum Occupancy 482

12.10.3 Communication Value 482
12.10.4 Spectrum Utilization 483
12.10.5 Spectrum Capacity 484
12.11 Conclusions 484
12.12 References 485

Appendix A. Atmospheric and Rain Data 487

Appendix B. PDF of a Signal with Interference and Noise 497

B.1 Introduction 497
B.2 References 500

Index 501

Preface

The growing demand for high-speed data connections to serve a variety of business and personal uses has driven an explosive growth in telecommunications technologies of all sorts including optical fiber, coaxial cable, twisted-pair telephone cables, and wireless. Nations have recognized that telecommunications infrastructure is as significant as roads, water systems, and electrical distribution in supporting economic growth. In developing countries it is not particularly unusual to see cell phone service in a town or village that does not yet have a water or sewer system. In the United States, recent government initiatives have recognized the importance of broadband telecommunications to economic growth.

This book focuses on fixed broadband wireless communications – a particular sector of the communication industry that holds great promise for delivering high-speed data to homes and businesses in a flexible and efficient way. The concept of 'broadband' communications is a relative one. Compared to the 1200-baud modems commonly used 20 years ago, today's dial-up phone connections with 56-kbps modems are 'broadband'. The demands and ambitions of the communication applications and their users have expanded, and will continue to expand, on what is meant by 'broadband'. The term is evolving, as is the technology that is classified as broadband. Nevertheless, for the purposes of this book I will use the somewhat arbitrary definition that broadband wireless systems are those designed for, and capable of handling baseband information data rates of 1 Mbps or higher, knowing that future developments may well move this threshold to 5 or 10 Mbps and beyond. The term 'broadband' also has an engineering significance that will be discussed in some detail in this book. Broadband wireless channels, as distinguished from narrowband channels, are those whose transfer characteristics must be dealt with in a particular way, depending on the information transmission speed and the physical characteristics of the environment where the service is deployed.

The term 'fixed' has also become somewhat nebulous with the technological developments of the past few years. Whereas fixed and mobile were previously well-understood differentiators for system types, we now have intermediate types of network terminals including fixed, portable, nomadic, and mobile, among others. Recent system standards such as those for 3G UMTS W-CDMA define different service levels and data rates depending on whether the user is in a fixed location, walking, or moving at high speed. This trend portends a convergence of fixed and mobile system types whose operation and availability are largely transparent to the application users. As will be shown, whether the system user is at a fixed location or in motion affects several decisions about the system design, the most appropriate technology, and the quality and performance that can be expected from a wireless application.

Although there have been a few books recently written on broadband, and specifically wireless broadband, in general they have been intended for non-technical audiences.

This book is intended for engineers who are faced with designing and deploying fixed broadband wireless systems, and who must also have sufficient understanding of the theory and principles on which the designs are based to formulate creative solutions to special engineering problems that they will eventually face. Along with generally accepted design assumptions and simplifications, the underlying theory and requisite mathematics are included where necessary to provide this foundation knowledge.

In addition to design engineers who deal with fixed broadband wireless systems on a daily basis, this book is also well suited to graduate and post-graduate level courses that are focused on wireless communications engineering. Wireless communication system design and planning is an increasingly important area that warrants serious academic treatment.

This book also covers some areas that have not classically fallen in the domain of wireless RF engineers; in particular, traffic modeling, environment databases, and mapping. Wireless system design is driven by the commercial requirements of the system operators who ultimately build viable businesses by successfully serving the traffic demands of the customers in their service areas. Detailed statistical modeling of packet-based traffic for a variety of applications (e-mail, web-browsing, voice, video streaming) is an essential consideration in fixed broadband system design if the operator's capacity and quality of service objectives are to be achieved.

The chapters in this book are organized with the fundamentals of electromagnetic propagation, channel and fading models, antenna systems, modulation, equalizers and coding treated first since they are the building blocks on which all wireless system designs are based. Chapters on multiple access methods and traffic modeling follow. The remaining chapters set forth the specific details of many types of line-of-sight (LOS) and non-line-of-sight (NLOS) systems, including elemental point-to-point links as well as point-to-multipoint, consecutive point, and mesh networks. Because of their importance, a separate chapter is devoted to designing both LOS and NLOS point-to-multipoint networks. The final chapter deals with the important subject of channel assignment strategies where the capacity and service quality of the wireless network is ultimately established.

Fixed wireless design relies on a number of published sources for data and algorithms. For convenience, the essential data, such as rain rate tables and maps, is included in the Appendices. In general, the referenced publications chosen throughout are currently available books or journal papers which are readily accessible in academic libraries or on-line. For the most recent or unique work, technical conference papers are also utilized.

A book of this type is clearly not a solo effort. I would like to thank several people who offered valuable comments, including Tim Wilkinson for reviewing Chapters 7 and 8, George Tsoulos for reviewing Chapter 6, and Jody Kirtner for reviewing Chapter 5, and for her efforts in proofreading the entire manuscript. Creating and refining a technical work such as this book is an evolutionary process where comments, suggestions, and corrections from those using it are most welcome and encouraged. I hope and anticipate that this book will prove to be a worthwhile addition to the engineering libraries of those who design, deploy, and manage fixed broadband wireless systems.

Harry R. Anderson
Eugene, Oregon, USA
January, 2003.

1

Fixed broadband wireless systems

1.1 INTRODUCTION

The theoretical origin of communications between two points using electromagnetic (EM) waves propagating through space can be traced to James Maxwell's treatise on electromagnetism, published in 1873, and later to the experimental laboratory work of Heinrich Hertz, who in 1888 produced the first radio wave communication. Following Hertz's developments at the end of the nineteenth century, several researchers in various countries were experimenting with controlled excitation and propagation of such waves. The first transmitters were of the 'spark-gap' type. A spark-gap transmitter essentially worked by producing a large energy impulse into a resonant antenna by way of a voltage spark across a gap. The resulting wave at the resonant frequency of the antenna would propagate in all directions with the intention that a corresponding signal current would be induced in the antenna apparatus of the desired receiving stations for detection there. Early researchers include Marconi, who while working in England in 1896 demonstrated communication across 16 km using a spark-gap transmitter, and Reginald Fassenden, who while working in the United States achieved the first modulated continuous wave transmission. The invention of the 'audion' by Lee DeForest in 1906 led to the development of the more robust and reliable vacuum tube. Vacuum tubes made possible the creation of powerful and efficient carrier wave oscillators that could be modulated to transmit with voice and music over wide areas. In the 1910s, transmitters and receivers using vacuum tubes ultimately replaced spark and arc transmitters that were difficult to modulate. Modulated carrier wave transmissions opened the door to the vast frequency-partitioned EM spectrum that is used today for wireless communications.

Radio communications differed from the predominate means of electrical communication, which at the time was the telegraph and fledgling telephone services. Because the new radio communications did not require a wire connection from the transmitter to the receiver as the telegraph and telephone services did, they were initially called *wireless communications*, a term that would continue in use in various parts of the world for several

Fixed Broadband Wireless System Design Harry R. Anderson
© 2003 John Wiley & Sons, Ltd ISBN: 0-470-84438-8

decades. The universal use of the term *wireless* rather than *radio* has now seen a marked resurgence to describe a wide variety of services in which communication technology using EM energy propagating through space is replacing traditional wired technologies.

1.2 EVOLUTION OF WIRELESS SYSTEMS

As the demand for new and different communication services increased, more radio spectrum space at higher frequencies was required. New services in the Very High Frequency (VHF) (30–300 MHz), Ultra High Frequency (UHF) (300–3,000 MHz), and Super High Frequency (SHF) (3–30 GHz) bands emerged. Table 1.1 shows the common international naming conventions for frequency bands. Propagation at these higher frequencies is dominated by different mechanisms as compared to propagation at lower frequencies. At low frequency (LF) and Mediumwave Frequency (MF), reliable communication is achieved via EM waves propagating along the earth–atmosphere boundary – the so-called *groundwaves*. At VHF and higher frequencies, groundwaves emanating from the transmitter still exist, of course, but their attenuation is so rapid that communication at useful distances is not possible. The dominant propagation mechanism at these frequencies is by space waves, or waves propagating through the atmosphere. One of the challenges to designing successful and reliable communication systems is accurately modeling this space-wave propagation and its effects on the performance of the system.

The systems that were developed through the twentieth century were designed to serve a variety of commercial and military uses. Wireless communication to ships at sea was one of the first applications as there was no other 'wired' way to accomplish this important task. World War I also saw the increasing use of the wireless for military communication. The 1920s saw wireless communications used for the general public with the establishment of the first licensed mediumwave broadcast station KDKA in East Pittsburgh, Pennsylvania, in the United States using amplitude modulation (AM) transmissions. The 1920s also saw the first use of land-based mobile communications by the police and fire departments where the urgent dispatch of personnel was required.

From that point the growth in commercial wireless communication was relentless. Mediumwave AM broadcasting was supplemented (and now largely supplanted) by

Table 1.1 Wireless frequency bands

Frequency band	Frequency range	Wavelength range
Extremely low frequency (ELF)	<3 kHz	>100, 000 m
Very low frequency (VLF)	3–30 kHz	100,000–10,000 m
Low frequency (LF)	30–300 kHz	10,000–1,000 m
Mediumwave frequency (MF)	300–3,000 kHz	1,000–100 m
High frequency (HF)	3–30 MHz	100–10 m
Very high frequency (VHF)	30–300 MHz	10–1.0 m
Ultra high frequency (UHF)	300–3,000 MHz	1.0–0.1 m
Super high frequency (SHF)	3–30 GHz	10–1.0 cm
Extra high frequency (EHF)	30–300 GHz	1.0–0.1 cm

frequency modulation (FM) broadcasting in the VHF band (88–108 MHz). Television appeared on the scene in demonstration form at the 1936 World Fair in New York and began widespread commercial deployment after World War II. Satellite communication began with the launch of the first Russian and American satellites in the late 1950s, ultimately followed by the extensive deployment of geostationary Earth orbit satellites that provide worldwide relay of wireless communications including voice, video, and data.

Perhaps the most apparent and ubiquitous form of wireless communication today are cellular telephones, which in the year 2002 are used by an estimated one billion people worldwide. The cellular phone concept was invented at Bell Labs in the United States in the late 1960s, with the first deployments of cell systems occurring in the late 1970s and early 1980s. The so-called third generation (3G) systems that can support both voice and data communications are now on the verge of being deployed.

Fixed wireless systems were originally designed to provide communication from one fixed-point terminal to another, often for the purpose of high reliability or secure communication. Such systems are commonly referred to as 'point-to-point (PTP)' systems. As technology improved over the decades, higher frequency bands could be successfully employed for fixed communications. Simple PTP telemetry systems to monitor electrical power and water distribution systems, for example, still use frequencies in the 150- and 450-MHz bands. Even early radio broadcast systems were fixed systems, with one terminal being the transmitting station using one or more large towers and the other terminal the receiver in the listener's home. Such a system could be regarded as a 'Point-to-Multipoint (PMP)' system. Similarly, modern-day television is a PMP system with a fixed transmitting station (by regulatory requirement) and fixed receive locations (in general). Television can also be regarded as 'broadband' using a 6-MHz channel bandwidth in the United States (and as much as 8 MHz in other parts of the world), which can support transmitted data rates of 20 Mbps or more.

The invention of the magnetron in the 1920s, the 'acorn' tube in the 1930s, the klystron in 1937, and the traveling wave tube (TWT) in 1943 made possible efficient ground and airborne radar, which saw widespread deployment during World War II. These devices made practical and accessible a vast new range of higher frequencies and greater bandwidths in the UHF and SHF bands. These frequencies were generically grouped together and called *microwaves* because of the short EM wavelength. The common band designations are shown in Table 1.2. Telephone engineers took advantage of the fact that

Table 1.2 Microwave frequency bands

Microwave band name	Frequency range (GHz)
L-band	1–2
S-band	2–4
C-band	4–8
X-band	8–12
Ku band	12–18
K-band	18–27
Ka band	27–40

PTP microwave links used in consecutive fashion could provide much lower signal loss and consequently higher quality communication than coaxial cables when spanning long distances. Although buried coaxial cables had been widely deployed for long-range transmission, the fixed microwave link proved to be less expensive and much easier to deploy. In 1951, AT&T completed the first transcontinental microwave system from New York to San Francisco using 107 hops of an average length of about 48 km [1]. The TD-2 equipment used in this system were multichannel radios manufactured by Western Electric operating on carrier frequencies of around 4 GHz. Multihop microwave systems for long-distance telephone systems soon connected the entire country and for many years represented the primary mechanism for long-distance telecommunication for both telephone voice and video. The higher frequencies meant that greater signal bandwidths were possible – microwave radio links carrying up to 1800 three-kilohertz voice channels and six-megahertz video channels were commonplace.

On the regulatory front, the Federal Communications Commission (FCC) recognized the value of microwave frequencies and accordingly established frequency bands and licensing procedures for fixed broadband wireless systems at 2, 4, and 11 GHz for common carrier operations. Allocations for other services such as private industrial radio, broadcast studio-transmitter links (STLs), utilities, transportation companies, and so on were also made in other microwave bands.

Today, these long-distance multihop microwave routes have largely been replaced by optical fiber, which provides much lower loss and much higher communication traffic capacity. Satellite communication also plays a role, although for two-way voice and video communication, optical fiber is a preferred routing since it does not suffer from the roughly 1/4 s round-trip time delay when relayed through a satellite in a geostationary orbit 35,700 km above the Earth's equator.

Today, frequencies up to 42 GHz are accessible using commonly available technology, with active and increasingly successful research being carried out at higher frequencies. The fixed broadband wireless systems discussed in this book operate at frequencies in this range. However, it is apparent from the foregoing discussion of wireless system evolution that new semiconductor and other microwave technology continues to expand the range at which commercially viable wireless communication hardware can be built and deployed. Frequencies up to 350 GHz are the subject of focused research and, to some extent, are being used for limited military and commercial deployments.

The term *wireless* has generally applied only to those systems using radio EM wavelengths below the infrared and visible light wavelengths that are several orders of magnitude shorter (frequencies several orders of magnitude higher). However, free space optic (FSO) systems using laser beams operating at wavelengths of 900 and 1100 nanometers have taken on a growing importance in the mix of technologies used for fixed broadband wireless communications. Accordingly, FSO systems will be covered in some detail in this book.

1.3 MODELS FOR WIRELESS SYSTEM DESIGN

The process of designing a fixed broadband wireless communications system inherently makes use of many, sometimes complex, calculations to predict how the system

will perform before it is actually built. These models may be based on highly accurate measurements, as in the case of the directional radiation patterns for the antennas used in the system, or on the sometimes imprecise prediction of the levels and other characteristics of the wireless signals as they arrive at a receiver. All numerical or mathematical models are intended to predict or simulate the system operation before the system is actually built. If the modeling process shows that the system performance is inadequate, then the design can be adjusted until the predicted performance meets the service objects (if possible). This design and modeling sequence make take several iterations and may continue after some or all of the system is built and deployed in an effort to further refine the system performance and respond to new and more widespread service requirements.

The ability to communicate from one point to another using EM waves propagating in a physical environment is fundamentally dependent on the transmission properties of that environment. How far a wireless signal travels before it becomes too weak to be useful is directly a function of the environment and the nature of the signal. Attempts to model these environmental properties are essential to being able to design reliable communication systems and adequate transmitting and receiving apparatus that will meet the service objectives of the system operator. Early radio communication used the LF portion of the radio spectrum, or the so-called long waves, in which the wavelength was several hundred meters and the propagation mechanism was primarily via groundwaves as mentioned earlier. Through theoretical investigation starting as early as 1907 [2], an understanding and a model of the propagation effects at these low frequencies was developed. The early propagation models simply predicted the electric field strength as a function of frequency, distance from the transmitter, and the physical characteristics (conductivity and permittivity) of the Earth along the path between the transmitter and receiver. The models themselves were embodied in equations or on graphs and charts showing attenuation of electric field strength versus distance. Such graphs are still used today to predict propagation at mediumwave frequencies (up to 3000 kHz), although computerized versions of the graphs and the associated calculation methods were developed some years ago [3].

All wireless communication systems can be modeled using a few basic blocks as shown in Figure 1.1. Communication starts with an information source that can be audio, video, e-mail, image files, or data in many forms. The transmitter converts the information into a signaling format (coding and modulation) and amplifies it to a power level that is needed to achieve successful reception at the receiver. The transmitting antenna converts the transmitter's power to EM waves that propagate in the directions determined by the design and orientation of the antenna. The propagation channel shown in Figure 1.1 is not a physical device but rather represents the attenuation, variations, and any other distortions that affect the EM waves as they propagate from the transmitting antenna to the receiving antenna.

By using EM waves in space as the transmission medium, the system is necessarily exposed to sources of interference and noise, which are often beyond the control of the system operator. Interference generally refers to identifiable man-made transmissions. Some systems such as cellular phone systems reuse frequencies in such a way that interference transmitters are within the same system and therefore can be controlled. Cellular system design is largely a process of balancing the ratio of signal and interference levels to achieve the best overall system performance.

7

Modulation, equalizers, and coding

7.1 INTRODUCTION

The preceding chapters have dealt with the mechanisms that result in transmitted electromagnetic signals arriving at a receiver and the modeling techniques that can be used to represent these mechanisms in the system design process. The quality of the wireless communication circuit depends on the physical propagation mechanisms and on the nature of the transmitted signal itself. Since communication engineers have complete control over how the transmitted signal is created and how it is detected, many of the advances in communications technology in recent years have been in the way signals are coded, modulated, demodulated, decoded, and otherwise structured to maximize efficient, multiple access use of the spectrum. Many of these developments have been specifically designed to provide more robust performance in the presence of the propagation channel variations and distortions described in earlier chapters.

The original point-to-point microwave systems that were used for long-distance, multihop telephone circuits were analog systems in which a large number (up to 1,800) of single-sideband (SSB) 3-kHz voice channels were multiplexed together into a single analog signal that was used to frequency-modulate (FM) a microwave radio carrier. Such analog systems dominated the microwave industry for a few decades but progressively were replaced by systems using digital multiplexing and modulation in the 1970s and 1980s. The ability to correct transmission errors using coding techniques and the ability to regenerate a noise-free signal at each microwave repeater made the digital radios a logical improvement.

This change was paralleled in the cellular radio industry. The first-generation cellular systems built in the early 1980s used analog voice signals to frequency-modulate a radio frequency (RF) carrier operating in channels 25 or 30 kHz wide. These were later replaced with second-generation digital systems using various modulation and multiple access

Fixed Broadband Wireless System Design Harry R. Anderson
© 2003 John Wiley & Sons, Ltd ISBN: 0-470-84438-8

methods. Even the radio and TV broadcast industries, which have used analog transmission for several decades, are now slowly transitioning to digital transmission systems. Although analog still plays a role when used by older incumbent fixed systems, new fixed broadband system deployments will undoubtedly be digital systems; so only digital modulation will be treated in this chapter.

Because modulation, equalizers, and coding are features of the system that can be completely controlled, the effort that has been devoted to this area in communication system design has been enormous, with an associated enormous volume of literature on the subject. There are several books available at both the beginner and the advanced level that deal with modulation, equalizers and coding. Several of these books are used as references here and listed at the end of the chapter. The information in this chapter is intended to extract and focus on those basic elements of modulation, equalizers, and coding that are germane to fixed broadband wireless system design. Specifically, this includes the performance of various technologies in the presence of noise and interference, and the ability of equalizers and coding to improve performance by mitigating the effects of noise, interference, and channel response impairments. For the system design process, the signal-to-interference + noise ratio ($S/(I+N)$ ratio) is a fundamental design objective. The $S/(I+N)$ ratio or signal-to-interference + noise ratio (SINR) objective is also a direct indicator of the quality and robustness of the signal received by the end user.

Modern digital communication systems including those for fixed broadband usually have adaptable modulation that responds to the conditions of the channel (also called *link adaptation*). Under nominally good channel conditions, the transmission will use the highest data rate available. As the channel deteriorates because of a rain fade or some other reason, the system will automatically revert to a more robust modulation and coding scheme with associated lower net data throughput to maintain the communication link. Adaptability creates a degree of complication for the system designer, which will be addressed in later chapters.

7.2 DIGITAL MODULATION – AMPLITUDE, FREQUENCY, AND PHASE

A transmitted signal has three fundamental characteristics – frequency, amplitude, and phase. These characteristics can be changed individually or in combination in response to the information that is to be transmitted. It is assumed that the information itself is already in digital form. For an information signal that is basically analog like speech, a digital-to-analog converter (DAC) is used along with some means of compressing the signal so that the fewest number of bits necessary are used to achieve communication with the desired fidelity. Regardless of the means used to create the information signal, it ultimately becomes a time sequence of binary digits ('bits') that can have only one of two states: 0 or 1.

Digital modulation is the process of modifying the frequency, phase, or amplitude of a signal in response to whether a 0 or a 1, or a string (block) of 0s and 1s, is to be sent. From the terminology of the earliest digital systems (telegraph and the radio Morse

code), this process is still called 'keying', so that fundamental modulation types can be grouped as:

- *ASK*: Amplitude Shift Keying in which the amplitude of the signal is changed among discrete amplitude levels to represent a pattern of bits. This includes the simplest form, on–off keying (OOK), in which the signal is turned on and off (from maximum amplitude to zero amplitude) in response to a bit pattern. OOK is used by free space optic systems.
- *FSK*: Frequency Shift Keying in which the frequency of the signal is changed among discrete frequencies to represent a pattern of bits.
- *PSK*: Phase Shift Keying in which the phase of the signal is changed among discrete phase states to represent a pattern of bits.

A wide variety of modulation methods are derived from these basic forms. As the number of available discrete amplitude levels, frequencies, or phase states increases, a greater number of more information bits can be conveyed with one signal state. For example, with a simple ASK system with two possible amplitudes, only one bit can be conveyed at a time – a 1 with the high amplitude and a 0 with the low amplitude. The voltage/frequency/phase that is transmitted to represent a bit pattern is called a *channel symbol* or simply *symbol*. The binary amplitude symbol in this case can represent one bit. If the symbol could have four amplitude levels instead of two, then each symbol could represent two bits (00, 01, 10, and 11). Similarly, as the number of possible states increases, the number of bits that can be conveyed increases such that

$$n = \log_2(M) \text{ or } M = 2^n \tag{7.1}$$

where M is the number of signal states a symbol can have and n is the number of information bits it can convey. This leads to the obvious observation that to increase the data rate or capacity of a link, symbols with a large number of signal states should be used. In a world with no noise, no interference, and no channel impairments, this would be the clear thing to do. The problem, of course, is that there is noise, interference, and channel impairments in a real system, which makes it increasingly difficult to correctly detect which signal state has been transmitted as the number of signal states increases (for a given transmitter power). Detection errors potentially lead to a higher bit error rate (BER) at the receiver and to a decrease in the ultimate signal quality delivered to the end user. The impact of noise and interference, and the associated trade-offs with modulation type, will be discussed in Section 7.4.

7.3 FIXED BROADBAND WIRELESS MODULATION METHODS

Given the broad flexibility that comes from being able to change the amplitude, frequency, or phase of the signal, there have been a large number of modulation schemes developed over the years. Of these, a few have emerged as the most attractive for fixed broadband

wireless use as well as general wireless communication such as cellular and other mobile radio applications. The discussion of the modulation types is confined to these types.

7.3.1 BPSK, QPSK, $\pi/4$-DQPSK

Two of the simplest and the most robust modulation methods are binary phase shift keying (BPSK) and quadrature phase shift keying (QPSK). As the names imply, BPSK switches between two phase states to convey the bit pattern, while QPSK switches between four phase states. The associated signal states or *signal constellations* for these two types are shown in Figures 7.1(a and b), respectively. BPSK can take on phase values of 0 and 180 degrees. QPSK can have phase values of 45, 135, 225, and 315 degrees. Since each phase state for QPSK conveys two bits of information, the data used with QPSK can be coded so that adjacent phases differ by just one bit as shown in Figure 7.1(b). Since most detection errors result in an adjacent phase being detected (not two phase states away), this encoding approach means that a symbol detection error will result in one bit error rather than two. This type of coding is called *Gray encoding* [1].

From Figure 7.1(b) it is clear that certain bit pattern transitions will cause the signal vector to go through zero (00 to 11, for example), which in turn will cause the carrier amplitude to go to zero. This transition through zero carrier amplitude can cause adjacent channel spectral components that result in adjacent channel interference, depending on the linearity of the power amplifier (PA) over this signal power transition range. To eliminate the phase transition through zero, a form of QPSK called $\pi/4$-DQPSK is used in which the 2-bit pattern is conveyed by phase transitions among eight-phase states (separated by an angle of $\pi/4$) rather than an absolute phase as depicted in Figure 7.1(b). Because the transmitted bits are detected by the phase difference from one symbol to the next, it is *differential* QPSK or DQPSK. The source bit sequence translates to the phase transitions shown in Table 7.1.

In addition to BPSK and QPSK, eight-phase modulation, or 8PSK, has grown in interest recently because of its choice of an enhanced cellular system with higher data rates called EDGE. With 8PSK, the phase states are separated by 45 degrees rather than 90 degrees as with QPSK. With 8 states, 3 bits of information are conveyed with each channel symbol.

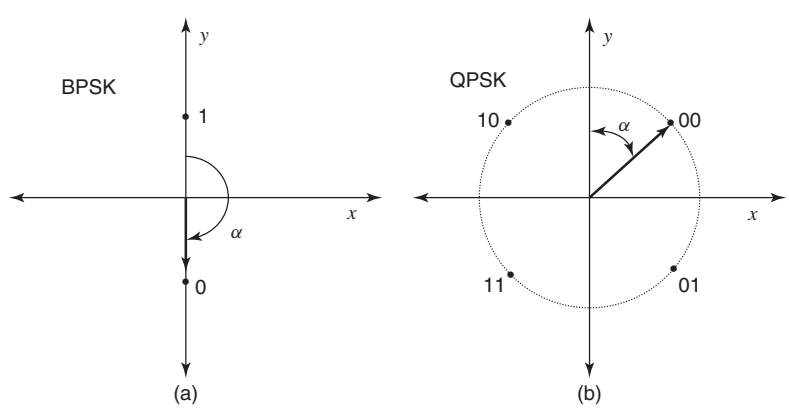

Figure 7.1 Signal constellations for (a) BPSK and (b) QPSK.

Table 7.1 $\pi/4$-DQPSK phase transitions

Bit pattern	Phase transition
00	$-3\pi/4$
01	$-\pi/4$
10	$+\pi/4$
11	$+3\pi/4$

The power spectra of BPSK and QPSK depend primarily on the symbol pulse shape that is used in the transmitter. There are a wide variety of choices here but the one most commonly used is the *raised cosine pulse*. The time domain representation of this pulse is given by

$$x(t) = \frac{\sin \pi t/T}{\pi t/T} \frac{\cos(\pi \alpha t/T)}{1 - 4\alpha^2 t^2/T} \tag{7.2}$$

where T is the symbol time length, α is a pulse shape or *rolloff* parameter ($0 \leq \alpha \leq 1.0$), and t is the time. The spectrum of this pulse is given by [2]

$$X(f) = \left\{ \begin{array}{ll} T & \text{for } 0 \leq |f| \leq (1-\alpha)/2T \\ \dfrac{T}{2}\left[1 + \cos\left(\dfrac{\pi T}{\alpha}\right)\left(|f| - \dfrac{1-\alpha}{2T}\right)\right] & \text{for } \dfrac{1-\alpha}{2T} \leq |f| \leq \dfrac{1+\alpha}{2T} \\ 0 & \text{for } |f| > (1+\alpha)/2T \end{array} \right\} \tag{7.3}$$

The time and frequency domains for the raised cosine pulse are drawn in Figures 7.2(a and b), respectively, for a few different values of α . Figure 7.2(a) shows that the complete time domain function is noncasual (starts before zero time). This is normally addressed by injecting a delay of 3 or 4 symbol times to permit a reasonably accurate representation of the leading tail of the pulse to be transmitted and the expected spectrum shape to be closely realized.

When this pulse shape is used with BPSK and QPSK, the resulting power spectral density (PSD) for the transmitted signal is given by [3]

$$S(f) = \frac{A^2}{T}|X(f)| \tag{7.4}$$

This is basically the same shape as Figure 7.2(b) except that it is scaled for the pulse width and amplitude. Note that the PSD for QPSK is the same as that for $\pi/4$-DQPSK.

The analysis of the error rates for BPSK and QPSK as a function of noise and interference follows in Section 7.4.

7.3.2 16QAM, 64QAM, and 256QAM

The number of bits per transmitted symbol with BPSK and QPSK is rather modest for a given bandwidth. To increase the data rate, symbols that convey more information

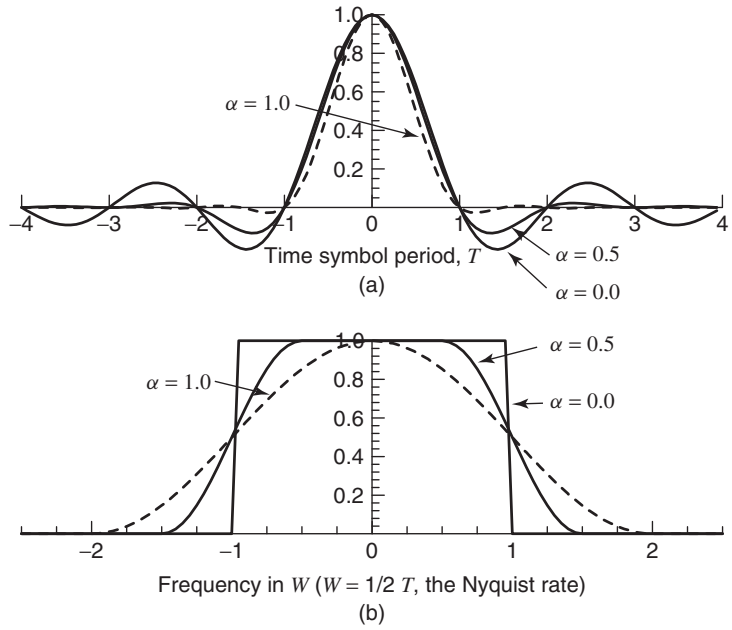

Figure 7.2 Raised cosine spectrum: (a) time domain and (b) frequency domain.

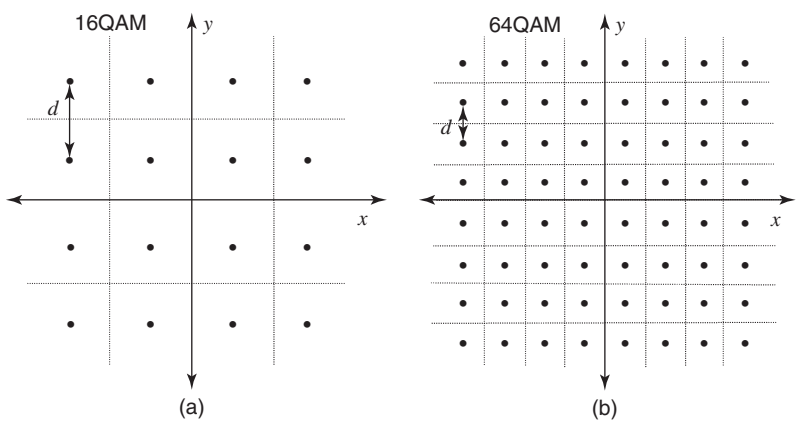

Figure 7.3 Signal constellations for (a) 16QAM and (b) 64QAM.

bits with more signal constellation states are required. The family of modulation types known as *quadrature amplitude modulation* (QAM) is the usual choice for achieving higher symbol efficiencies. Figure 7.3 illustrates the signal constellations for 16QAM and 64QAM. 256QAM is the same as 64QAM except with a 16×16 square array of constellation states.

From the constellations it is apparent that the QAM family of modulations are closely related and in fact can be generated using the same basic in-phase/quadrature (I/Q) modulator circuitry. Comparing Figures 7.1(b) and 7.3(a), it is also evident that QPSK is a member of this family that could be called 4QAM and is generated using the same I/Q modulator. With these four symmetric constellations – 4QAM, 16QAM, 64QAM, and 256QAM, the symbol transmission efficiencies ranging from 2 bits/symbol to 8 bits/symbol are possible. Other constellations such as 32QAM and 128QAM are also possible and sometimes used, but the four symmetric constellations described here are the most common because of the ease with which they are generated and because of their ability to be readily transitioned to higher or lower efficiency QAM constellations.

The PSD for QAM modulation is the PSD of the transmitted symbol provided the transmitted symbols are uncorrelated; that is, the autocorrelation function of the input data sequence is 0 for $\tau > T/2$ and $\tau < -T/2$ where τ is the time offset variable. If the raised cosine pulse described in the last section is used for the symbol shape, the spectrum shown in Figure 7.2(b) applies to QAM.

The analysis of the error rates for the QAM modulation as a function of the noise and the interference levels follows in Section 7.4.

7.3.3 Orthogonal frequency division multiplexing (OFDM)

Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier modulation method that uses a number of carriers (called subcarriers), each operating with a relatively low data rate. Taken together, the composite data of all the subcarriers is comparable to the data rate possible using the same basic modulation at a higher data rate with a single carrier in the same channel bandwidth. The main advantage of OFDM is that the symbol duration can be much longer so that the susceptibility to errors from intersymbol interference (ISI) due to multipath time dispersion is greatly reduced. From a frequency domain perspective, a frequency-selective fade will cause a problematic fading depth on only a few subcarriers. Errors will potentially occur on the few bits associated with those subcarriers but not the others, so the net error rate for all subcarriers taken together can still be made acceptably low if coding is also employed.

If a block of N_s symbols is considered, in OFDM each is used to modulate a subcarrier, so there are also N_s subcarriers. If each symbol has duration T_s, the time domain signal representation for the OFDM during T_s can be written as

$$s(t) = \sum_{N_1}^{N_2} d_{i+N_s/2} \exp\left[j2\pi \frac{n}{T_s}(t - t_s) \right] \text{ for } t_s \le t \le t_s + T_s \tag{7.5}$$

where $N_1 = N_s/2$, $N_2 = N_s/2 - 1$, and t_s is a symbol start time. Since the amplitudes and phases of the subcarriers in the transmitted spectrum $S(f)$ represent the N_s data symbols, the time domain in (7.5) is actually the inverse Fourier transform (IFFT) of the data block. As discussed in [4], the transform can be performed by an inverse fast Fourier transform (IFFT) or inverse discrete Fourier transform (IDFT), the choice being driven simply by which of these is the most computationally efficient.

It can be observed from (7.5) that this is a sum of N_s quadrature cosine and sine functions with frequencies $2\pi n/T_s$. These frequencies are integer multiples of each other so that they are inherently orthogonal, or

$$\int_{t_s}^{t_s+T_s} \exp(j2\pi nt/T_s)\exp(j2\pi mt/T_s)\,dt = 0 \quad n \neq m \tag{7.6}$$

$$\int_{t_s}^{t_s+T_s} \exp(j2\pi nt/T_s)\exp(j2\pi mt/T_s)\,dt = T_s \quad n = m \tag{7.7}$$

where n and m are integers. The receiver detector has the same subcarrier frequencies that are synchronized with the transmitted carriers. The demodulated OFDM signal using the orthogonal subcarriers recovers the original data symbols as implied by (7.7). The N_s detected data symbols can then be reassembled into the data block.

The number of subcarriers is determined by the total required data rate and the maximum delay spread that will be experienced in the channel. The individual subcarriers can be modulated using PSK or QAM as described in the preceding sections. The error performance of the separate subcarriers can be considered independently as long as the subcarriers remain orthogonal and there is no intercarrier interference.

At the receiver the amplitude and phase distortions introduced by the channel must still be equalized. In the case of differential QPSK, only the phase distortions in the channel are relevant to detection and they may be corrected using a differential equalizer that considerably simplifies the receiver [5]. The use of QAM or other amplitude-dependent constellations on a subcarrier requires amplitude equalization as well, so there is strong motivation to use BPSK, QPSK, or other phase-only constellations. The use of amplitude-dependent modulation like QAM also aggravates the peak-to-average power ratio problem discussed in Section 7.3.3.1.

For channels with deep fades at some frequencies, the subcarriers at these faded frequencies will experience very high error rates because of low signal-to-noise ratio (SNR) values. The error performance on these weak channels can dominate the overall error rate, thus setting a lower bound on the error performance of OFDM. Therefore, for reasonable error rate performance, it is essential for OFDM systems to have coding and interleaving. Coded OFDM is known as COFDM. Coding and interleaving are discussed in Section 7.6.

OFDM can also be sensitive to carrier frequency errors that degrade the orthogonality between subcarriers. This is known as intercarrier interference (ICI). This requirement can impose carrier accuracy demands on OFDM systems that are several orders of magnitude higher than those for single-carrier systems of similar capacity [6].

To ensure that time dispersion in the form of multipath does not introduce intersymbol interference (ISI – see Section 7.4.4), a guard interval is usually added into the system that extends the symbol duration. This guard interval is actually a cyclic extension of the subcarrier wave. The guard time will be determined by the delay spread of the channel that the system is expected to accommodate. Since the guard interval is not used to transmit information, it directly reduces the data capacity of the OFDM link.

The power spectrum of an OFDM system is the composite of the power spectra of the individual subcarriers. As the number of subcarriers increases, the spectrum approaches the

rectangular band-limited spectrum of ideal Nyquist filtering. This is one of the advantages of OFDM in applications in which the spectrum is densely occupied and in which controlling adjacent channel interference is important. An example of an OFDM spectrum for 16 subcarriers is shown in Figure 7.4 for a rectangular pulse amplitude shape for the transmitted symbol ($\alpha = 0.0$) and with a raised cosine pulse shape with $\alpha = 0.1$. The out-of-band spectrum can be further attenuated using the raised cosine pulse, but this erodes the orthogonality of the subcarriers, which in turn degrades the received error rate [3]. Note that the raised cosine function is applied to the time domain pulse, and not the frequency domain as shown in Figure 7.2.

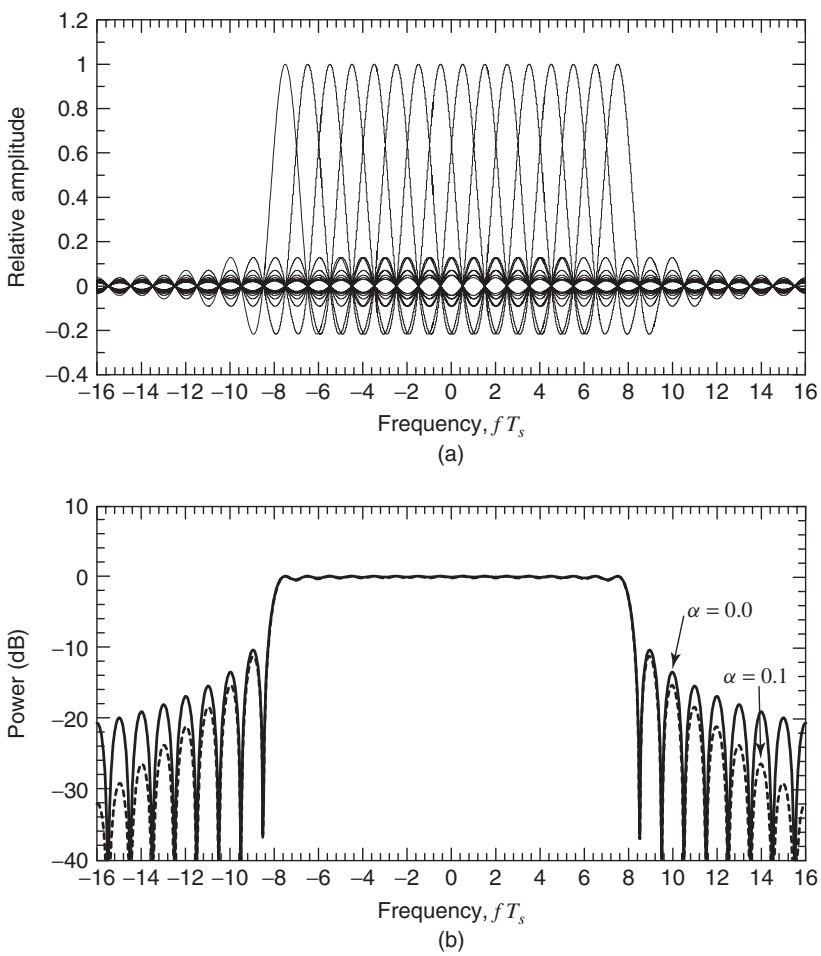

Figure 7.4 Spectrum for OFDM with $N = 16$. (a) Overlapping subcarrier spectra with rectangular pulse and (b) composite power spectrum with rectangular pulse and raised cosine pulse with $\alpha = 0.1$.

7.3.3.1 OFDM peak-to-average power ratio

Depending on the data signal being transmitted, it is possible that several of the subcarriers could be at their maximum values during a symbol time. If the sum of an unlimited number of (uncorrelated) subcarriers is considered, the probability density function (pdf) of the resulting signal envelope is a Rayleigh distribution with the probability of high peak values diminishing as the value increases. The peak power needed to represent this summation without distortion could be considerably higher than the average power of the waveform, greatly increasing the maximum power and linearity requirements on the transmitter power amplifier (PA).

A number of approaches for dealing with this problem have been explored over the past several years. The simplest is to clip the signal amplitude in a clipping 'window', which is what the PA ultimately does anyway, but this produces distortion and out-of-band emissions that are undesirable. As discussed in [7], several variations on the windowing approach show various levels of out-of-band emission. As shown in [8], it is possible to apply a code to the transmitted data block that essentially prevents subcarriers from adding simultaneously to limit the peak to a certain value above the average power level. A suitable code set turns out to be complementary Golay codes [9]. The coding method has a relatively small coding overhead (a 3/4 rate code can achieve a peak-to-average ratio (PAR) of 4 dB) and is currently the preferred approach.

These various methods have not eliminated the PAR problem but have reduced it to a manageable level. OFDM has thus become a viable and important option for fixed broadband wireless systems, especially for non-line-of-sight (NLOS) point-to-multipoint systems.

7.4 ERROR PERFORMANCE WITH NOISE AND INTERFERENCE

A fundamental measure of the capability of a digital modulation scheme is its BER in the presence of noise and interference. The BER is a function of the ratio of the desired signal level to the noise and interference levels, decreasing as the ratio decreases. In the system design process, accurate predictions of the S/N ratio (SNR) and the S/I ratio (signal-to-interference ratio (SIR)) are sought so that the link BER can be accurately predicted and the overall link quality assessed.

In considering the impact of interference and noise on signal detection errors, the underlying objective is to develop system design criteria that can be used to determine whether the outage performance of a particular link will meet the service objectives of the system operator. The error performance of different modulation types degrade differently with noise and interference. The higher-level modulation type where the constellation states are closer together will obviously have more detection errors at lower noise and interference levels than low-level modulations with more widely spaced constellation states.

For practical system design, the required SNR and SIR levels for a given maximum tolerable error rate threshold is usually given by the equipment manufacturer's specifications. These values can normally be employed directly in the link design. However, it is

useful to understand the fundamental way in which errors in digital systems are caused by noise and interference so that trade-offs between high-level, high-efficiency modulation with higher error rates versus low-level, low-efficiency modulation with lower error rates can be recognized. The purpose of this section is to provide this information using a fundamental description of the error-creating processes that occur in digital signals.

7.4.1 Error performance with Gaussian noise only

The BER for BPSK and QPSK with noise can be found in many books on digital communications. Regardless of the modulation type, the same basic signal geometry illustrated in Figure 7.5 is applicable. The received signal vector $s(t)$ is perturbed by Gaussian noise vectors in quadrature creating a resultant vector $r(t)$ that is presented to the detector. Since n_x and n_y are jointly Gaussian random variables with $\sigma_x^2 = \sigma_y^2 = N_0/2$, then n_x and $s + n_y$ are also jointly Gaussian with $\sigma_r^2 = N_0/2$. The joint pdf is then

$$f(n_x, s + n_y) = \frac{1}{2\pi\sigma_r^2} \exp\left[-\frac{(n_y^2 + n_x^2)}{2\sigma_r^2}\right] \tag{7.8}$$

Since the phase $\theta = \tan^{-1}(n_x/(s + n_y))$ and the amplitude $r = \sqrt{n_x^2 + (s + n_y)^2}$, these variables can be used in (7.8) to get

$$p_{r,\theta}(r, \theta) = \frac{1}{2\pi\sigma_r^2} \exp\left(-\frac{r^2 + s^2 - 2sr\cos\theta}{2\sigma_r^2}\right) \tag{7.9}$$

Equation (7.9) can be integrated over r from 0 to ∞ to get $p_\theta(\theta)$, or over 0 to 2π to get $p_r(r)$. For PSK modulation, the phase pdf $p_\theta(\theta)$ is of interest. The result of this integration is given by [10]

$$p_\theta(\theta) = \int_0^\infty p_{r,\theta}(r, \theta)\, dr \tag{7.10}$$

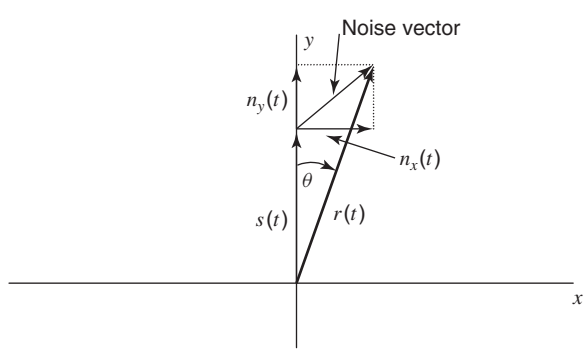

Figure 7.5 Signal vector with additive Gaussian noise.

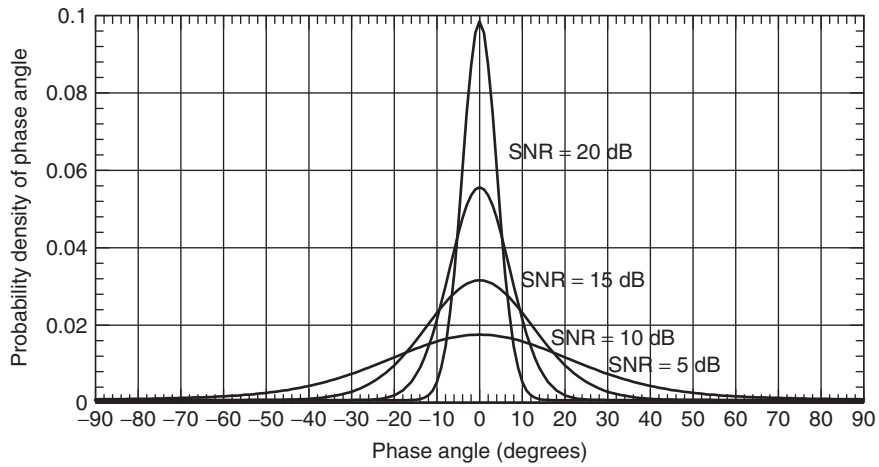

Figure 7.6 Probability of phase angle for PSK with Gaussian noise.

This is usually reduced to the form

$$p(\theta) = \frac{1}{2\pi} e^{-\gamma} \left(1 + \sqrt{4\pi\gamma} \cos\theta e^{\gamma \cos^2\theta} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2\gamma}\cos\theta} e^{-r^2/2}\, dr \right) \tag{7.11}$$

If the symbol SNR is defined as $\gamma = S^2/N_0$, (7.12) can be evaluated numerically to find the probability densities. Figure 7.6 shows $p_\theta(\theta)$ for several values of γ in dB. The probability of an error due to noise is found as the area under the pdf curve for the phase angles outside the range of $\pm\pi/M$ around the transmitted phase, where M is the number of phase states (2 for BPSK, 4 for QPSK, 8 for 8PSK, etc.). The integral in (7.10) only reduces to a simple form for BPSK and QPSK – for other cases it can be readily evaluated numerically. The assessment of the phase error assumes that a perfect, noise-free reference is available for coherent detection. For differential DBPSK and DQPSK, the transmitted symbol is determined by the phase difference between successive symbols. This allows for a simpler receiver but slightly degrades error performance since the phase detection reference is also corrupted by noise.

The formulas for the symbol error rate for BPSK, QPSK, and *M-ary* QAM, and a few other modulation types as a function of the ratio of the signal energy per bit to the noise power per bit (E_b/N_0 or γ_b) are shown in Table 7.2, along with the source reference for each formula. The equations in Table 7.2 give symbol error rates; the bit error rates can be found assuming Gray encoding using

$$P_b \cong \frac{1}{k} P_M \text{ where } k = \log_2 M \tag{7.12}$$

Equation (7.12) is most accurate for high values of SNR because it assumes that a symbol error will result in an adjacent symbol being detected, which differs from the correct

Table 7.2 Symbol error probability for various modulation types based on γ_b

Modulation type	Symbol error rate equation	Source
Coherent BPSK	$P_M = P_b = \frac{1}{2}\text{erfc}(\sqrt{\gamma_b})$	Reference [1] eqn. (4.2.106)
Differential BPSK	$P_M = P_b = \frac{1}{2}\exp(-\gamma_b)$	Reference [1] eqn. (4.2.117)
Coherent QPSK	$P_M = \text{erfc}(\sqrt{\gamma_b})[1 - \frac{1}{4}\text{erfc}(\sqrt{\gamma_b})]$	Reference [1] eqn. (4.2.107)
Differential QPSK	$P_M = 2[Q(a,b) - \frac{1}{2}I_0(ab)\exp[-\frac{1}{2}(a^2+b^2)]]$ $a = \sqrt{2\gamma_b(1-1/\sqrt{2})} \quad b = \sqrt{2\gamma_b(1+1/\sqrt{2})}$	Reference [1] eqn. (4.2.118)
Coherent M-ary PSK	$P_M \cong \text{erfc}\left(\sqrt{k\gamma_b}\sin\frac{\pi}{M}\right)\gamma_b \gg 1$	Reference [1] eqn. (4.2.109) and eqn. (4.2.110)
Differential M-ary PSK	$P_M \cong \sqrt{\frac{1+\cos(\pi/M)}{2\cos(\pi/M)}}\text{erfc}\sqrt{k\gamma_b[1-\cos(\pi/M)]}$ for $M \geq 3$	Reference [11] eqn. (46)
Coherent M-ary QAM	$P_M = 2\left(1-\frac{1}{\sqrt{M}}\right)\text{erfc}\left(\sqrt{\frac{3k}{2(M-1)}\gamma_b}\right) \times\left[1-\frac{1}{2}\left(1-\frac{1}{\sqrt{M}}\right)\text{erfc}\left(\sqrt{\frac{3k}{2(M-1)}\gamma_b}\right)\right]$	Reference [1] eqn. (4.2.144)
GMSK	$P_M = \frac{1}{2}\text{erfc}(\sqrt{\alpha\gamma_b}) \quad \alpha = 0.68$ for typical filter	Reference [12]

Note: erfc is the complementary error function; I_0 is the modified Bessel function of the first kind and order 0; $Q(a,b)$ is the Marcum Q function; M is the number of signal constellation states; $k = \log_2 M$; $\gamma_m = k\gamma_b$.

symbol by just one bit. With low SNR, the erroneously detected symbol may be farther away and two or more bits will be incorrect. The symbol error rates calculated from the formulas in Table 7.2 are plotted in Figure 7.7 for BPSK and QPSK and Figure 7.8 for QPSK, 16QAM, 64QAM, and 256QAM.

For a typical fixed wireless system, the amplitude of the signal will vary because of either multipath fading or rain fading events as described in Chapter 4. Since the noise power level from the receiving equipment is constant, the SNR will vary as the signal level fades. The relevant parameter, then, is the *instantaneous SNR* rather than the average SNR. The error rate then becomes a variable that follows the fading. When the signal level fades to a low level, the error rate rises during that fade, potentially resulting in an error burst that can destroy a block of data, or system synchronization, and timing. Considering that the 'bursty' nature of errors in these systems leads to the design concept of *fade margin* in wireless systems – the faded signal level that can occur before the error rate reaches an unacceptable threshold level. Fade margin is one of the most significant parameters in fixed broadband system design since it directly leads to a determination of the link availability. Fade margin will be employed in the link design discussion in later chapters.

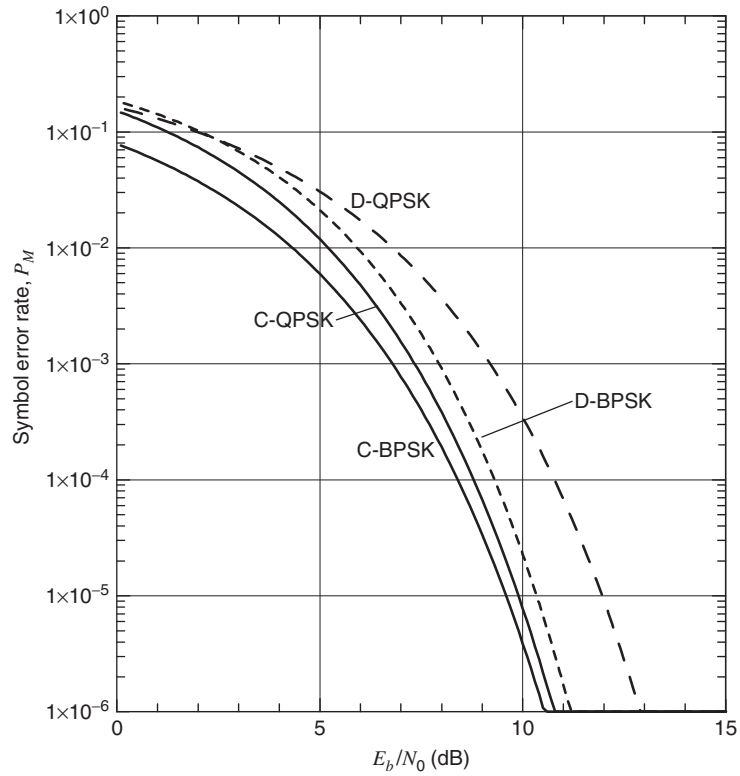

Figure 7.7 Probability of a symbol error for BPSK and QPSK.

The error rate performance shown in Figures 7.7 and 7.8 is the theoretical best-case performance for the indicated modulation type. For actual equipment, the error rate performance will usually be worse than the theoretical values due to a variety of nonideal aspects of the equipment design and operation. For example, the published specifications for a typical Local Multipoint Distribution Service (LMDS) system show an error performance for 16QAM of 10^{-6} for an SNR of 17 dB, 2 dB worse than the theoretical value. The theoretical curves therefore are best used as relative indicators of error rate performance and associated capacity. For link design, of course, the actual performance numbers of the equipment that will be deployed should be used. This is a parameter that can potentially differentiate one vendor from another.

7.4.2 Error performance with noise and constant amplitude interference

The vector approach used above can be extended to include a single sine wave interference vector as shown in Figure 7.9. The quadrature noise vectors are shown as n_x and n_y as before with the sine wave interference as a vector with amplitude b. The objective is to

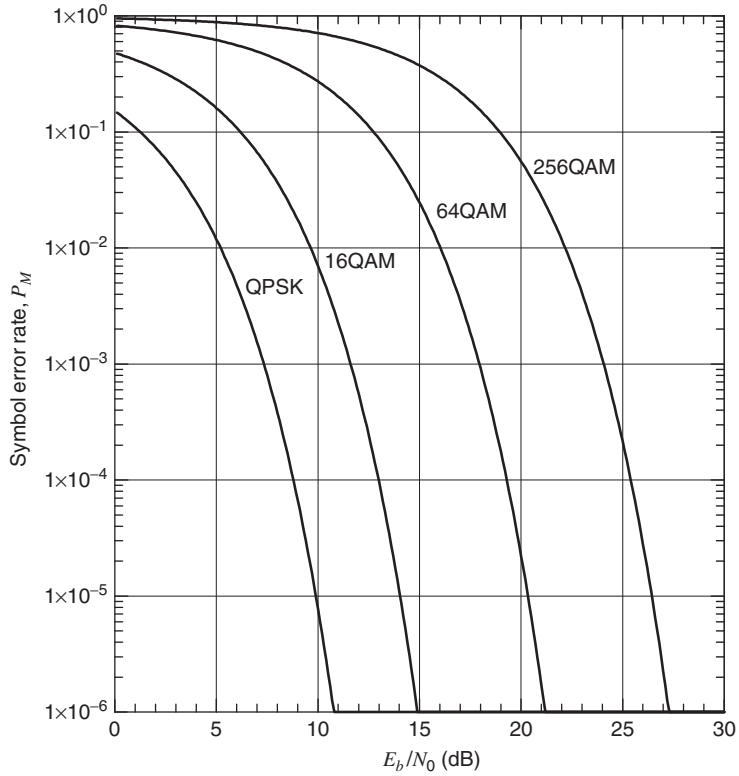

Figure 7.8 Probability of a symbol error for M-ary QAM.

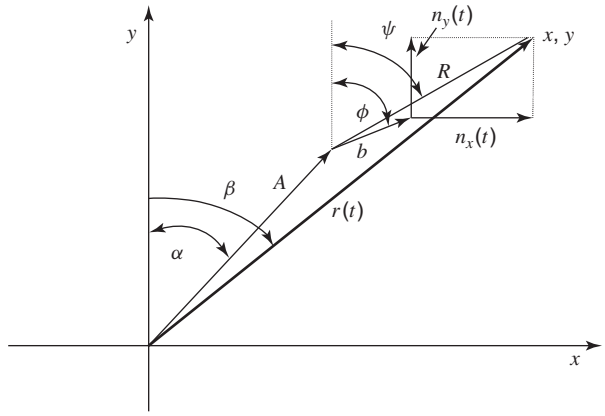

Figure 7.9 Signal, interference, and noise vectors for noise and sine wave interference case.

find the pdf of the phase and amplitude of this ensemble of vectors. Because an analysis of the pdf and error rate of digital modulation with a sine wave interferer and noise is rarely found in texts on digital communications, the details of the derivation of the pdf of x, y are included in Appendix B. The resulting pdf from Appendix B is

$$p(x, y \mid A, b) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x - A\sin\alpha)^2 + (y - A\cos\alpha)^2 + b^2}{2\sigma^2}\right] \times I_0\left(\frac{b\sqrt{(x - A\sin\alpha)^2 + (y - A\cos\alpha)^2}}{\sigma^2}\right) \tag{7.13}$$

where $I_0(\cdot)$ is the Bessel function of the first kind and of the order zero. To find the pdf of x or y alone, it is necessary to integrate (7.13) over one or the other random variables:

$$p(x) = \int_{-\infty}^{\infty} p(x, y \mid A, b)\, dy \quad p(y) = \int_{-\infty}^{\infty} p(x, y \mid A, b)\, dx \tag{7.14}$$

For mobile communications, the interference signal usually includes fading as does the desired signal, so a statistical description is normally used to describe the interference amplitude and the SNR. The aggregation of multiple random phase interferers, even if with constant amplitudes, also leads to a statistical description of the interference vector amplitude.

For fixed line-of-sight (LOS) systems, however, a single dominant interference signal can be received that exhibits very little fading due to high-gain antenna rejection of multipath components. The error rate analysis with a quasi-constant amplitude interferer is relevant to such systems.

The pdf in (7.13) does not easily yield a closed-form expression for the probability of error. However, it can readily be integrated by numerical methods to find the volume of the pdf in the error region for any particular modulation scheme, especially those in which the error boundaries between modulation states are rectilinear in x and y. As an example, this pdf will now be applied to finding error rates for 16QAM and QPSK when perturbed by noise and interference.

7.4.2.1 16QAM with noise and interference

As mentioned, a commonly used rectangular signal constellation in fixed wireless is 16QAM. The modulation constellation states are shown in Figure 7.3(a) along with the error boundaries (shown as dotted lines and the x-, y-axis). Coherent detection of *M-ary* QAM generally involves finding the Euclidean distance in the 2D signal space from the receive signal point to each of the M expected constellation states. The expected signal state that is closest to the received signal point is selected as the signal state that was transmitted. An error occurs when this distance is shorter to a signal state that is different from the transmitted signal state. The probability of error can be found by finding the volume of the pdf in (7.15), which crosses over the error boundaries in Figure 7.3(a) for a certain transmitted state.

With many modulation types such as 4QAM and *M-ary* PSK, the error boundaries relative to a signal state are the same for all of the M possible modulation states. Evaluating the probability of error P_e, or error rate, can then be done by considering the error for one signal state. However, for 16QAM it is clear from Figure 7.3(a) that the error boundaries are not the same for all signal states. There are actually three different types to be considered. First are the four corner points that only have error boundaries on two sides. Next are the eight 'edge' points with error boundaries on three sides. Finally, there are four interior points with error boundaries on four sides. The integration of the error region for each of these three types of signal states must be done with different integration limits that take into account the different error boundaries.

An approach to this integration that has simple limits is to integrate the volume under the 2D pdf inside the region where a given state has no error, and then to subtract that result from 1.0 to arrive at the error probability. However, with numerical integrations involving small numbers, this can be a difficult approach to make-work successfully. Instead, the straightforward approach of integrating the error region directly is used here, with the integration beginning in the low value 'tails' and progressing toward the higher values near the center of the pdf to avoid losing the contribution of these small values to the sum due to the round-off error.

The awkward integration limits can be resolved by choosing an integration region that easily encompasses all the relevant areas of the pdf in (7.13), both those where errors occur and those where no errors occur. In the case under study here the integration limits are set at $\pm 3d$. When the integration loops are sequenced through the values of x and y, each x, y pair is first evaluated to determine if it falls inside the 'no-error' limits. If it does, evaluating $p(x, y)$ and its contribution to the integration sum is skipped. This 'cut out' approach to integrating a complicated region works very effectively with the 16QAM case. It can also be used successfully when the error boundaries are not rectilinear as in the case of M-PSK modulation signal constellations and their derivative forms. The cut out region can even be defined by complicated polygons if necessary.

The numerical integration using the cut out approach can be written in the following form:

$$P_e = \sum_{-3d}^{3d} \sum_{-3d}^{3d} p(x, y) \Delta x \Delta y, \text{ subject to integration limits below} \tag{7.15}$$

The cut out integration limits the evaluation of (7.15) so that the results of the double summation are not included under the following conditions:

$$\begin{aligned} &x > -d/2, y > -d/2 \text{ for the corner signal states} \\ &-d/2 > x > d/2, y > -d/2 \text{ for the edge signal states} \\ &-d/2 > x > d/2, -d/2 > y > d/2 \text{ for the interior signal states} \end{aligned} \tag{7.16}$$

With these integration limits, the probability of a symbol error was calculated for 16QAM with 6 different levels of sine wave interference (signal-to-interference ratios or SIR) and a range of signal-to-noise ratios (SNR per bit). The results are shown in Figure 7.10.

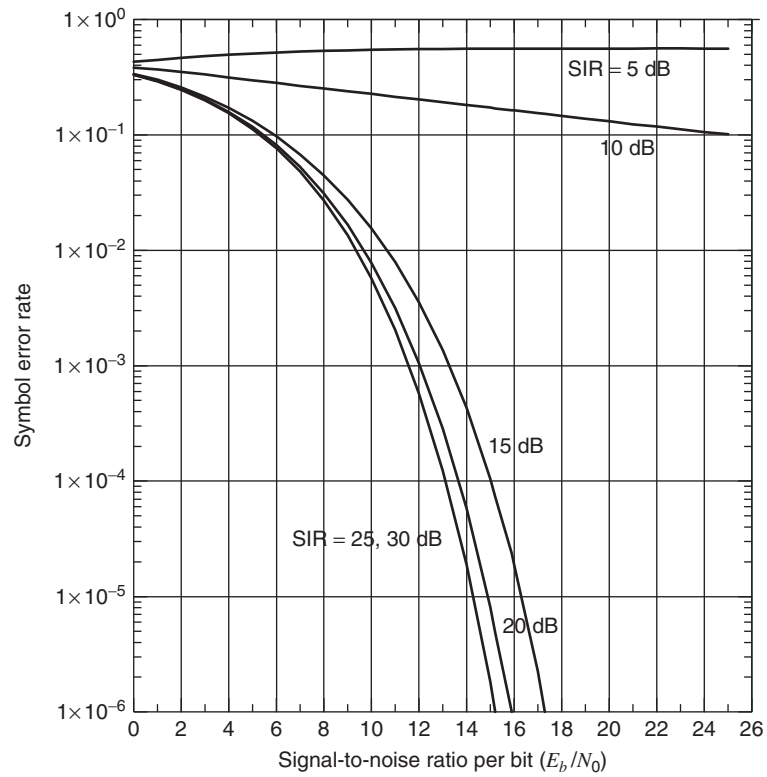

Figure 7.10 16QAM error rates with noise and constant amplitude interference.

The signal levels are the average signal levels for the 16 constellation states. The quantities SNR and SIR are defined as follows:

$$\mathrm{SNR} = 20\log\left(\frac{\overline{V}}{\sqrt{2}\sigma}\right)\ \mathrm{dB} \quad \mathrm{SIR} = 20\log\left(\frac{b}{\overline{V}}\right)\ \mathrm{dB} \tag{7.17}$$

where $\overline{V}$ is the average voltage of the signal modulation constellation. These values of SNR and SIR as defined in (7.17) will be used in the balance of this analysis.

As expected, with an SIR of 30 dB (essentially no interference) the error rate results are the same as the noise-only case and the curve in Figure 7.10 is in close agreement with [13]. As the interference level increases, the error rate also increases. At high levels of interference, a large portion of the pdf from (7.13) actually lies outside the no-error region even with no noise. At SIR = 5 dB, the error rate actually increases with *reduced* noise – an extraordinary condition in which the signal and interference pdf lies outside the interior signal state error boundaries such that only by adding noise does the signal occasionally fall within the 'no-error' region.

7.4.2.2 16QAM with 16QAM interference

The 16QAM error rate results for the constant amplitude interference can be applied to the problem of interference from another 16QAM signal in the same network. From Figure 7.3(a), it is clear that the interference vector b can take on one of the three discrete values, each with different probabilities. These are

- $b = d\sqrt{2}/2$ with probability 0.25
- $b = 3d\sqrt{2}/2$ with probability 0.25
- $b = d\sqrt{10}/2$ with probability 0.50

The error probability analysis described above can be repeated for each of these values of b, scaled by the overall S/I ratio. A similar approach can be applied to any single interferer with discrete amplitude values.

7.4.2.3 Coherent QPSK with noise and interference

The 2D pdf in (7.13) is given in terms of x and y, which is well suited to rectangular signal constellations like 16QAM. For angle constellations such as *M-ary* PSK, transforming (7.13) into polar coordinates is a more useful representation of the probability density. A typical example is coherent QPSK with a signal state constellation shown in Figure 7.1(b). With coherent detection it is assumed that some noise and interference-free phase reference is available against which the phase of the noise and interference-corrupted received signal can be evaluated.

The transformation from the rectangular coordinates of (7.13) to polar coordinates can be done by making the substitutions:

$$x = r\sin\beta \quad y = r\cos\beta \tag{7.18}$$

With these substitutions, (7.13) becomes

$$p(r,\beta \mid A, b) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(r\sin\beta - A\sin\alpha)^2 + (r\cos\beta - A\cos\alpha)^2 + b^2}{2\sigma^2}\right] \times I_0\left(\frac{b\sqrt{(r\sin\beta - A\sin\alpha)^2 + (r\cos\beta - A\cos\alpha)^2}}{\sigma^2}\right) \tag{7.19}$$

Recognizing that $dx\,dy = r\,dr\,d\beta$, the pdf for the phase can be found by integrating over the range of r:

$$p(\beta) = \int_0^\infty p(r,\beta) r\,dr \tag{7.20}$$

Several examples of $p(\beta)$ are shown in Figures 7.11(a and b) for different values of SNR and SIR. The pdf in (7.13) can also be used to create visually instructive 3-D plots of the overall signal constellation pdf. An example of such a plot for coherent QPSK with the indicated values of SNR and SIR is shown in Figure 7.12. The 'rings' in the pdf

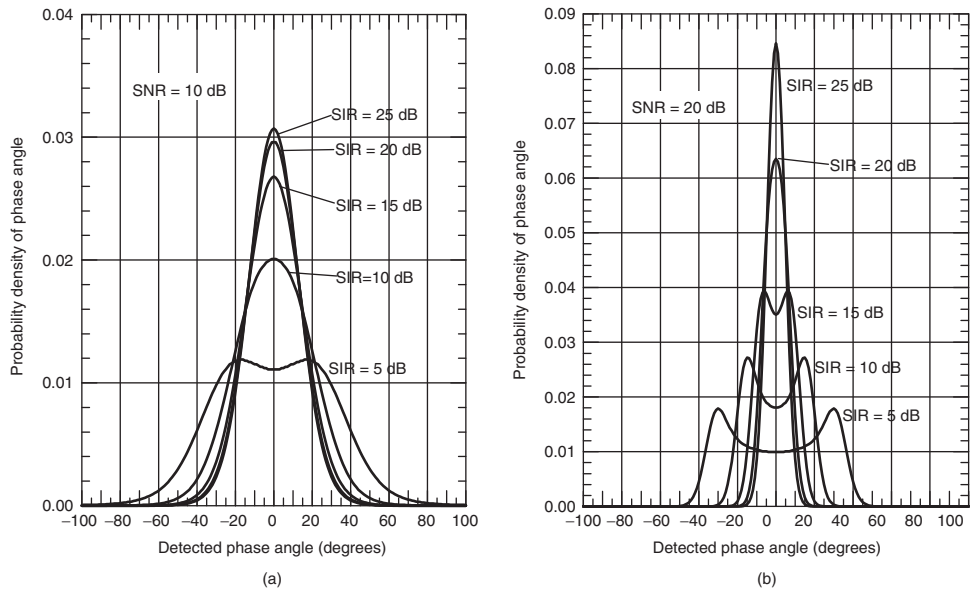

Figure 7.11 Phase pdf's for PSK with noise and single constant amplitude interference.

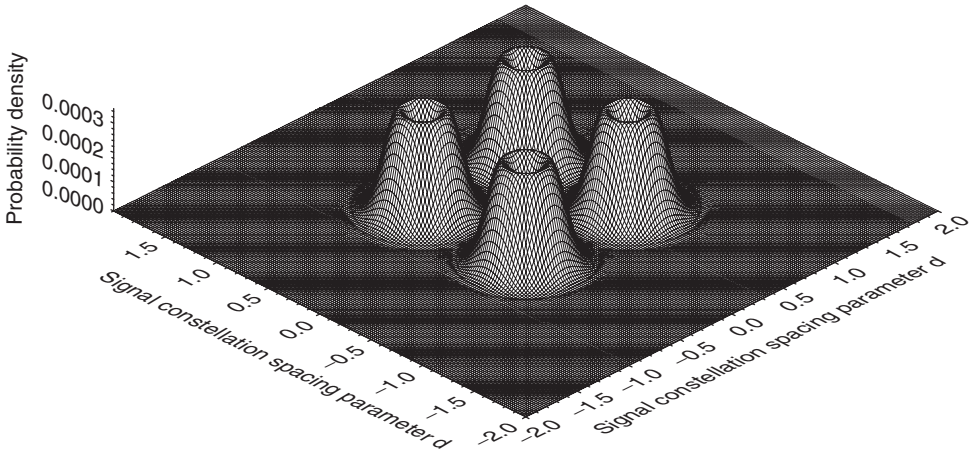

Figure 7.12 3D illustration of complete QPSK signal pdf with additive noise and single constant amplitude interference. SNR = 15 dB and SIR = 10 dB.

around each of the transmitted signal constellation points are clear. A similar 3D graph for 16QAM and an SIR value of 5 dB shows these ring pdf's crossing the error boundaries leading to the increasing error rates with lower noise as previously discussed.

Equation (7.20) can be evaluated numerically at a number of phase angles to provide a pdf description of the phase alone. The error rate P_e for an *M-ary* PSK system can then

be found by integrating $p(\beta)$ over the error region as follows:

$$P_e = \int_{-\pi}^{-\pi/M} p(\beta)\, d\beta + \int_{\pi/M}^{\pi} p(\beta)\, d\beta \tag{7.21}$$

Considering the symmetry of $p(\beta)$, (7.21) can be simplified to

$$P_e = 2\int_{\pi/M}^{\pi} p(\beta)\, d\beta \tag{7.22}$$

Equation (7.22) can now be applied directly to *M-ary* PSK system. Using (7.22), the symbol error rate for coherent QPSK was computed for the same range of SIR and SNR values as that used for the 16QAM results in Figure 7.10. The results for QPSK are shown in Figure 7.13. As expected, the performance of QPSK with equivalent noise and interference is superior to that of 16QAM (ignoring the different capacity of the two modulation types). This is especially true for low values of SIR where QPSK has not yet reached the error region even with an SIR value of 5 dB.

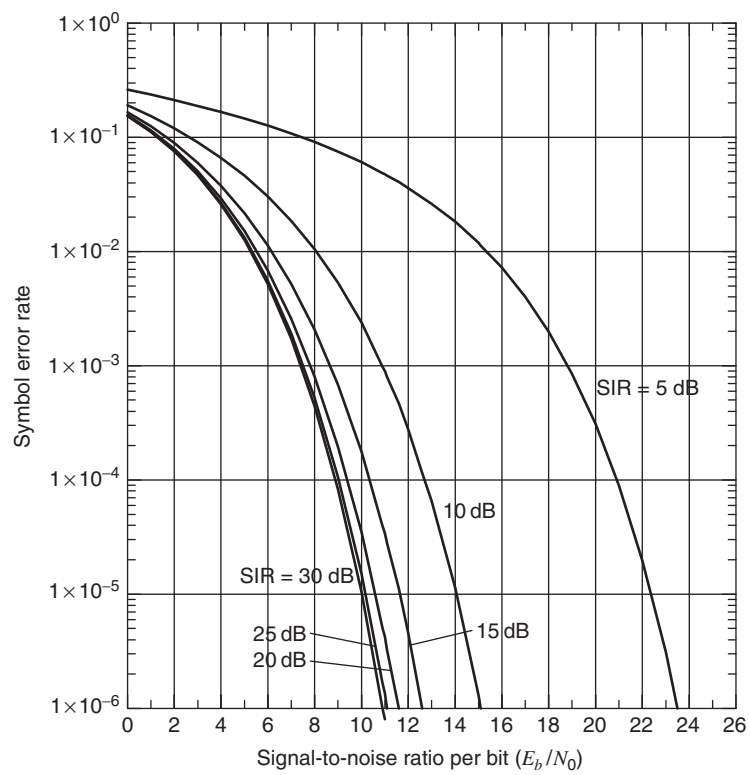

Figure 7.13 Coherent QPSK error rates with noise and constant amplitude interference.

The error rate results for QPSK were found using the pdf in (7.20). However, for this particular constellation they could have also been found using the joint pdf for x and y in (7.13). This approach would correspond to using only the corner cut out integration limits in (7.16).

7.4.2.4 Differential QPSK with noise and interference

The previous examples for predicting error rates in the presence of single-carrier interference were developed for coherent detection of 16QAM and QPSK, where a noise-free and an interference-free phase and amplitude reference was available. Many system architectures have been devised to recover a suitably uncorrupted phase and amplitude reference for use in wireless channels.

However, important classes of modulation schemes that offer somewhat simpler hardware implementation are differential techniques in which the information data is conveyed by the transition from one constellation state to another. The receiver then detects the difference between successive received symbols and interprets the differences in terms of the encoding scheme used at the transmitter. A commonly used example mentioned before is differential QPSK, or DQPSK, in which the change in phase, rather than the phase relative to a reference, is used to convey data.

The general technique described above for analyzing BER in the presence of noise and interference for coherent detection can be extended to the differential case by assuming that signal vectors for successive symbols are similarly perturbed by noise and interference as was shown in Figure 7.9. This approach was taken in [14] when it was assumed that the interference phase would undergo a fixed (but unknown) phase shift θ from symbol to symbol. An upper bound error rate result was then developed for DQPSK as a function of θ. The analysis can proceed in a manner similar to that given above with a resulting pdf for the phase *difference* between successive symbols. The error rate can then be determined by integrating the area under the pdf, which is outside the correct symbol detection bounds as before.

7.4.3 Error performance with flat-fading signal and interference

The analysis of the errors in the preceding section were done for discrete values of SNR and SIR as shown by the curves in Figures 7.10 and 7.13. As explained above, for short-range LOS links, this model may be applicable to the noise and interference circumstances of real links. For NLOS links, however, the desired signal will experience fading so that the SNR and SIR vary in the same way as the signal level variations assuming that the interference power is constant. Incorporating the pdf of the desired signal fading variations into the pdf for the overall signal amplitude given in (7.13) results in

$$p(x, y \mid b) = \int_0^\infty p(A)\, p(x, y \mid A, b)\, dA \tag{7.23}$$

Similarly, the fading distribution of the interference can be incorporated as

$$p(x, y) = \int_0^\infty p(b)\, p(x, y \mid b)\, db \tag{7.24}$$

Since there are no known closed-form solutions for (7.23) or (7.24), as a practical matter the error probability that results from the numerical integration of (7.13) for various values of A and b provides a tabular description that can be used as a lookup table in evaluating link error rates.

For example, Figure 7.10 shows the error rates for values of SNR and SIR. If the desired signal fades to a certain level with a given probability, the resulting SNR will occur with that probability so that the associated error rate will also occur with that probability. The probability of this error rate occurring can then be compared against the system performance criteria. The fading distribution for the interference can be handled in a similar numerical fashion. The degree of independence of the variations in SIR and SNR is important for fixed broadband systems, especially for correlated rain fading on desired and interfering paths as discussed in Section 4.4.

7.4.3.1 Noise approximation of interference

An accurate continuous or discrete (tabular) function for $p(A)$ or $p(b)$ is unlikely to be available for most systems, so the common approach is to make assumptions about the distribution of the desired and the interfering signal levels. This approach is discussed in the next section. The most common and straightforward approach to handling interference in fixed LOS broadband systems is to simply regard the interference power to be additional noise power. The instantaneous SNR $\rho(t)$ is

$$\rho(t) = \frac{s(t)}{N'(t)} \tag{7.25}$$

where the power of $N'(t)$ is the noise power plus the power of each of the interferers

$$N' = N + \sum_{j=1}^{J} I_j \tag{7.26}$$

The validity of this assumption improves as the number of interferers increases. Even with as few as six sine waves with random phases, the envelope begins to approach a Rayleigh distribution like noise [15]. When the interference amplitude is also varying, by the Central Limit Theorem, the probability distribution of the sum will approach a Gaussian distribution with a variance equal to the sum of the variances (power) of the individual interferers. This approximation also represents a conservative approach that results in higher error probabilities than that would occur if the interferers were treated with their real fading distributions. For real transmitters, the interfering powers are always necessarily limited in amplitude (a Gaussian distribution has no amplitude limits). For the occasion in which the interference is known to come from a single identifiable interferer, the explicit pdf approach detailed above will provide more accurate error probability predictions.

7.4.4 Error performance with frequency selective signal fading

The preceding sections discussed error rate performance in terms of noise and external independent interference sources. For a propagation environment that creates multiple

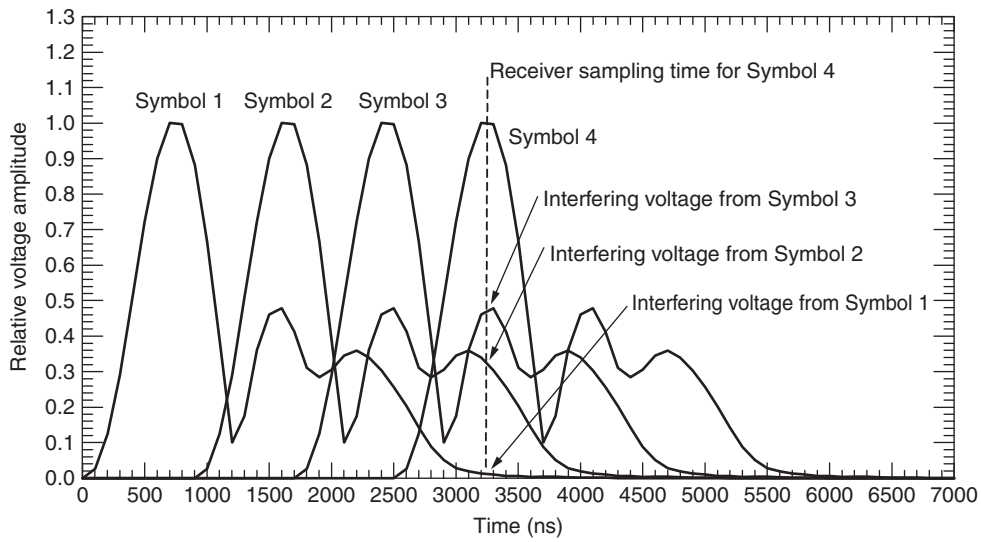

Figure 7.14 Intersymbol interference for a channel with multipath.

signal paths between the transmitter and the receiver, there is an additional interference mechanism due to components of the delayed multipath signals upsetting the proper detection of some subsequently transmitted symbol. This situation is illustrated in Figure 7.14. This is known as intersymbol interference or ISI.

The signal at the sampling instant can be modeled using the interference vector in Figure 7.9. The probability density for the composite signal vector developed in Section 7.4.2 can also be applied to find the error rate due to ISI. Of course, the multipath signals are typically not going to be static but vary with some distribution that can be postulated as Rayleigh or Rician (see Chapter 4).

The real microwave systems (RMS) delay spread (see Section 3.3.1) is often used as an indicator of the error rate for such time-dispersive channels. From the illustration in Figure 7.14, it is straightforward to envision two delay profiles each with the same value of RMS delay spread, but one with a low-amplitude multipath vector with a large delay, and the second with a high-amplitude multipath vector with a small delay. The high-amplitude, short-delay channel will have a higher error rate than the channel with the low-amplitude, long-delay multipath. RMS delay spread can therefore be a very a misleading indicator of error rate if no equalization or exploitation of the multipath is used in the receiver. For simple receivers without equalization, much more specific information about the nature of the multipath (amplitude, delay, variability) is needed before error rates can be found for such channels.

Unlike errors resulting from noise and external interference, the error rate resulting from the ISI cannot be reduced by increasing the transmitter power because this will simply increase the multipath signal amplitudes by the same proportional amount; for example, the SIR of the ISI stays the same. Instead, ISI can be effectively combated using equalizers of various designs as discussed in Section 7.5.

Multipath is not necessarily detrimental to system performance; multipath can also be exploited with two techniques:

- Using a Rake receiver, it is possible to reduce fading and improve error performance, essentially achieve time-diversity through combining of the signals on the branches of the Rake receiver. Rake receivers are discussed in Section 8.4.4.1 of Chapter 8.
- A multipath time-dispersive channel is indicative of an uncorrelated scattering channel that can be exploited to increase link capacity through the use of multiple-input, multiple-output (MIMO) technology, or increase link robustness with space-time codes (STCs). MIMO techniques are discussed in Section 6.7; STCs are discussed in Section 7.6.7.

A receiver must have sufficient bandwidth to take advantage of this additional decorrelated received signal energy provided via the multipath signals. For systems with limited bandwidth, time dispersion will manifest itself as narrowband frequency flat-fading for which alternate remedies such as diversity reception and pilot-assisted gain control are available to improve error performance.

7.5 EQUALIZERS

For systems in which multipath is expected and error rates must be maintained below an acceptable threshold level, there are equalizer solutions that essentially remove most or all the symbol distortion resulting from the multipath energy before the symbol detection is done. Two general types of equalizers are discussed here.

7.5.1.1 Time domain symbol equalizers

Time domain equalizers have been employed in digital point-to-point microwave link equipment for some time. In Section 4.2.4.2 of Chapter 4, the quantity 'dispersive fade margin' was defined along with the standard procedure for measuring it. The dispersive fade margin is directly related to the performance of the receiver equalizer and its ability to correct frequency-selective channel distortions. Microwave radios without equalizers typically have dispersive fade margins of 35 dB. Adding an effective equalizer can raise the dispersive fade margin to 55 or 60 dB.

A linear transversal equalizer has time-delayed signal taps as shown in Figure 7.15. The signals from the taps are multiplied by coefficients that are independently adjusted such that the multipath components for a received symbol are adjusted to a low level and the ISI is effectively removed. Within this general framework, there are several strategies for adjusting the taps. The earliest and simplest is the *zero-forcing* (ZF) equalizer [16] in which the signal additions due to the multipath are subtracted out by properly adjusting the tap coefficients c_j in Figure 7.15, forcing the ISI to zero. The equalizer uses a training symbol of known value to initially adjust the equalizer tap coefficients. If the channel is not varying, these tap settings can suffice. If the channel does vary as it normally does with wireless communications, the coefficients need to be periodically updated. Another

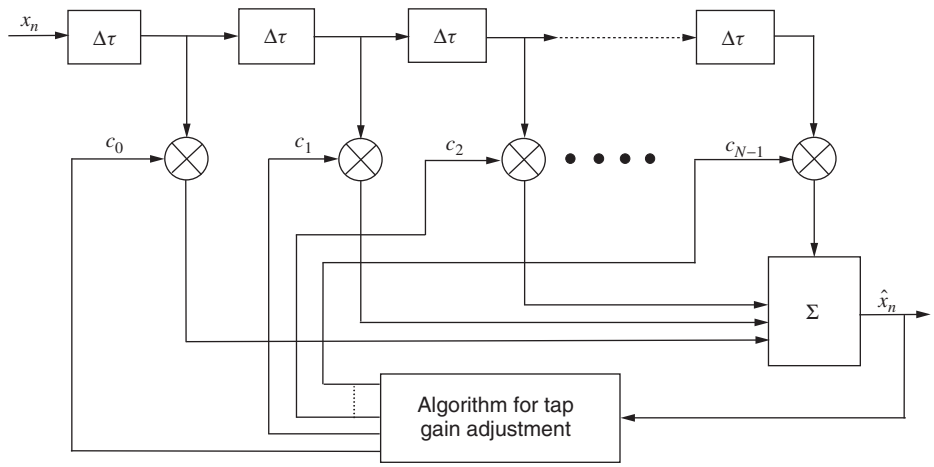

Figure 7.15 Linear transversal equalizer.

approach is the minimum mean-square error (MMSE) equalizer [17] in which the strategy is to minimize the mean-square error between the received signal, including residual ISI and noise, and the known training symbol.

The tap delay spacings in the original adaptive equalizers were set at the symbol transmission interval T. Since the multipath components can arrive at arbitrary times depending on the nature of the propagation environment, the time resolution of the equalizer adjustments was not ideal and resulted in high sensitivity to choosing the sampling time. This led to the development of *fractionally spaced equalizers* (FSE) in which the tap delay spacing was less than T. As described in [17], the FSE equalizer can perform as well as a symbol-spaced equalizer that is preceded by a matched filter. Since the matched filter is hard to realize in practice when the channel response is not known or varying, the FSE represents the more attractive solution.

Linear equalizers have the drawback of enhancing noise through the multiplicative process in the tap coefficients, especially when the channel has severe amplitude distortion leading to some high tap gains. This deficiency led to the development of the *decision feedback equalizer* (DFE), also called a nonlinear equalizer. A DFE uses previous detection decisions to eliminate the ISI caused by previously detected symbols on the symbol currently being detected. Figure 7.16 is a block diagram of a DFE consisting of two equalizers – a feedforward filter and a feedback filter. The feedforward filter is generally a fractionally spaced linear equalizer as described above. The feedback equalizer is a symbol-spaced equalizer whose input is the set of previously detected symbols. One drawback with DFE's is error propagation, which occurs because of the feedback of the detection decision from one symbol to the following symbol.

For all adaptive equalizers, the tap coefficients are periodically set to achieve some objective such as zero ISI or MMSE. Although the tap settings can be found by solving a set of linear equations, typically a search algorithm is used to find the optimum settings. There are several types of search algorithms, among them the *steepest descent method* and

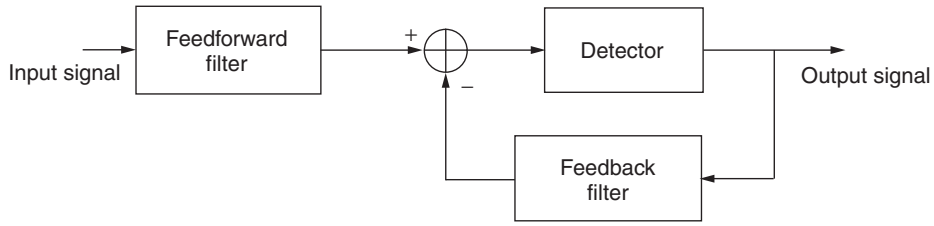

Figure 7.16 Nonlinear decision feedback equalizer DFE.

the *stochastic gradient algorithm,* also called the *least-mean-square* (*LMS*) *algorithm*. The intent of each is to rapidly find the coefficients with the least computation complexity. The LMS algorithm is the most popular approach because of its simplicity and stability [17].

These algorithms are iterative procedures in which an optimum solution can be approached arbitrarily closely depending on the number of iterations. Because each iteration takes one symbol time, and several hundred iterations may be required, the algorithm and gradient interval of the equalizer may be selected to be best suited for a particular type of wireless system. For example, with fixed broadband systems in which the channel response is reasonably assumed to be less variable than the channel response for a mobile radio system, the parameters of the adaptive algorithm for fixed broadband receivers can be adjusted to perform more efficiently in a more stable channel.

7.5.1.2 Frequency domain equalizers (FDE)

The objective of a frequency domain equalizer (FDE) is fundamentally the same as the time domain equalizer – to eliminate the ISI by correcting the channel response distortions. A frequency domain equalizer attempts to do this by equalizing the amplitude and phase of the channel response over the bandwidth of interest using a process that operates on a frequency domain representation of the signal. The concept of frequency domain equalization has actually been around for several decades; in fact, it is commonly found in consumer audio equipment at all levels in which the listener can manually adjust the amplitude or volume in various frequency bands to make up for frequency response deficiencies in the speaker system.

Ironically, adaptive FDE's were not considered a serious option for wireless communications until an OFDM system (Eureka-147) was deployed for digital audio broadcasting (DAB) in Europe in the 1980s. With the acceptance of the relative complexity of OFDM in a consumer system, especially the required fast Fourier transform (FFT's), the use of an FDE on a single-carrier system became not only plausible but potentially a more attractive alternative to the OFDM. For channels with long time delay ISI, FDE's are more effective and less complex than a DFE, to achieve the same net result.

A basic FDE is shown in Figure 7.17. The signal samples presented to the FFT can consist of one or more samples per received symbol. The M frequency domain samples from the FFT are multiplied by complex coefficients to correct the amplitude and phase of the frequency response. As with the time domain equalizer, a training symbol or sequence

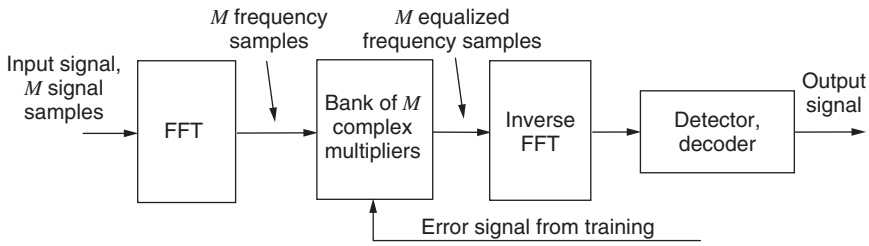

Figure 7.17 Frequency domain equalizer (FDE).

and similar steepest decent or LMS algorithms are used to find the optimum coefficients. Once equalized, the frequency domain samples are transformed back to the time domain using an inverse FFT where detection and decoding can be done.

As shown in [6], single-carrier QPSK with FDE (SC-FDE) can outperform uncoded OFDM in terms of error rate performance on both static and time-varying frequency-selective channels. This occurs because an FDE exploits the frequency diversity of the channel; OFDM requires coding to take similar advantage of frequency diversity. In [5], results were presented comparing SC-FDE to COFDM. With convolutional coding, interleaving, and weighted soft decisions, OFDM can produce error rate results that are comparable to SC-FDE in both static and time-varying frequency-selective channels. Further information on the application of FDE's to single-carrier modulation can be found in [18].

7.6 CODING TECHNIQUES AND OVERHEAD

The basic idea of coding is to add controlled redundancy to the transmitted signal that allows errors in data reception to be detected and corrected using forward error correction (FEC). Coding and FEC can detect and correct errors due to noise, fading, multipath, external interference, and other channel impairments. Codes can be designed that automatically correct a certain number of errors and detect an even larger number. For errors that are detected but not corrected, the receiver can request a retransmission of the data block that contained an error that could not be corrected. This is known as automatic retransmission request (ARQ) and is a necessary part of any wireless (or wired) data communications system. Of course, any retransmission detracts from the total link data throughput, so it is worthwhile to design a system to avoid such retransmissions. ARQ also causes a round-trip delay in the delivery of the data, which may make it unsuitable for real-time applications such as voice or video conversations. The usual emphasis, then, is to design coding schemes for use in wireless equipment that are sufficiently robust to accommodate both real-time and non real-time applications.

Any coding scheme necessarily adds 'overhead' to the signal. Overhead are data bits that the system needs for successful performance that do not carry information. Overhead also includes transmitted bits used for synchronization, equalizer training, system control and management, and so on. The bits that are exclusively used for information are

sometimes referred to as the *payload*. The system overhead may vary from 10 to 50% or more, a fact that should be borne in mind when the total link and net link capacities are considered with respect to the data rate delivery expectations of the end user.

Current coding schemes fall into one of the three general categories: block codes, convolutional codes, and space-time codes. These coding methods are discussed in the following sections.

7.6.1 Block codes

A block encoder takes a sequence or block of k data symbols and encodes it into a sequence or block of n code symbols. The encoded block is a vector called a *code word*; the number of redundant symbols is $n - k$, and the code is referred to as an (n, k) code. The degree of redundancy introduced into the code word by this process is called the *code rate*, R_c, and is given by the ratio k/n where $n > k$. The inverse of the code rate also indicates the amount of overhead added to the transmission by the coding. A rate $^3/_4$ code requires 4 code word symbols for every 3 data symbols, so the coding overhead is approximately 33%.

The elements in the code word are drawn from an alphabet of q elements. If the alphabet has only two members, 0 and 1, the code is a binary code and the elements in the code word are bits. If the number of elements is greater than 2 ($q > 2$), the code is nonbinary. It is clear that if a code word has n elements, the number of possible code words is q^n. For a binary code, the number of possible code words is 2^n, the number of code words needed to represent the possible information bit sequences is 2^k. The encoder will map a set of k information bits into an n length code word selected from the 2^k code words.

In addition to codes being classified as binary or nonbinary, they can also be classified as linear or nonlinear. If two code words C_i and C_j are considered, and a_1, a_2 are two elements of the alphabet, a code is said to be linear only if $a_1C_i + a_2C_j$ is also a code word. This also indicates that a linear code must contain an all-zero code word.

An additional basic parameter that describes a code is the *code word weight*. The code word weight is the number of nonzero elements that it contains. Considering the weights of all the codes is called the weight distribution or weight enumeration.

The basic objective in adding the redundant $n - k$ symbols is to increase the *Hamming distance* between any two code words. The Hamming distance is calculated as the number of symbols that are different between two code words. If the binary code word C_i has the sequence 00110011 and binary code word C_j has the sequence 01010101, the Hamming distance between C_i and C_j is 4. For a linear code that necessarily has a zero code word, the weight of the code word is the Hamming distance to the zero code.

For a minimum Hamming distance of $d_{\min}$ between any two code words, the code can correct t errors where t is given by

$$t \leq \text{int}\left(\frac{d_{\min} - 1}{2}\right) \tag{7.27}$$

The function int(·) is the integer function that rounds the argument down to the nearest integer value (since it is not possible to correct a fraction of a symbol). The minimum

Hamming distance $d_{\min}$ depends on the number of redundant code symbols $n - k$. The upper bound of the Hamming distance is given by $d_{\min} \leq (n - k) + 1$. If 4 redundant data symbols are added, the largest $d_{\min}$ can be is 5, and the largest number of bits in a code word that can be corrected from (7.27) is 2.

A number of binary linear codes have been developed over the years that differ primarily in how the code words are constructed. A *Hamming code* is one in which

$$(n, k) = (2^m - 1, 2^m - 1 - m) \tag{7.28}$$

where $m = n - k$ is any positive integer ($m \geq 2$) and $d_{\min} = 3$. From (7.28), with this minimum distance a Hamming code is capable of correcting one bit error, a rather limited capability compared to other coding schemes.

A *Hadamard code* is one in which the code words are selected from the rows of an $N \times N$ Hadamard matrix. A Hadamard matrix is a matrix of 0s and 1s in which any row differs from any other row in exactly $n/2$ vector elements. This necessarily implies that N must be an even integer. As a linear code, one row of the matrix contains all zeros and the other rows contain $n/2$ ones and $n/2$ zeros.

7.6.1.1 Cyclic codes

Cyclic codes are a type of linear block code that have defined relationships between code words that makes encoding and decoding easier to implement. A cyclic code is one in which code word $\mathbf{C}_i = [c_{n-1}, c_{n-2}, c_{n-3}, \ldots, c_0]$ can be formed into code word $\mathbf{C}_j = [c_{n-2}, c_{n-3}, c_{n-4}, \ldots, c_0, c_{n-1}]$ by a cyclic shift of the code elements. All cyclic shifts of $\mathbf{C}$ are therefore valid code words. Generating cyclic codes follows a straightforward algebraic process using generator polynomials. A detailed description can be found in [19].

A *Golay code* is a special (23,12) binary linear code with $d_{\min} = 7$. By adding a parity bit, it becomes an extended (24,12) Golay code with $d_{\min} = 8$. A Golay code can be generated as a cyclic code using a generator polynomial given in [19].

The Bose–Chaudhuri–Hocquenghem (BCH) codes are a large group of cyclic codes that can use both binary and nonbinary alphabets. Binary BCH codes have the parameters

$$\begin{aligned} n &= 2^m - 1 \\ n - k &\leq mt \\ d_{\min} &= 2t + 1 \end{aligned} \tag{7.29}$$

These parameters allow a wide range of codes with different code rates and error correction capabilities to be created. As with other cyclic codes, there are specific generator polynomials to create the codes. The coefficients of the generator polynomial for various block lengths are tabulated in various texts on digital communications and coding such as the one in [19].

As mentioned, BCH can be binary or nonbinary. An important subset of nonbinary BCH codes is *Reed–Solomon* (RS) codes. For nonbinary codes, the alphabet consists of

q elements ($q > 2$). As before, the length of a nonbinary code is indicated by N; the number of information symbols is indicted by K. The minimum distance is given by $D_{\min}$. The nonbinary block code is thus specified as (N, K) with $N - K$ added code bits. If the code is a *systematic code*, the code symbols are appended on to the end of the information symbols in the form of parity check symbols.

RS codes are described by the following parameters:

$$\begin{aligned} N &= q - 1 = 2^K - 1 \\ K &= 1, 2, 3, \ldots, N - 1 \\ D_{\min} &= N - K + 1 \\ R_c &= K/N \end{aligned} \tag{7.30}$$

The number of symbol errors, t, that the RS codes can correct has an upper bound of

$$\begin{aligned} t &= (D_{\min} - 1)/2 \\ &= (N - K)/2 \end{aligned} \tag{7.31}$$

RS codes are attractive for wireless communications because of the good distance properties between code words; in fact, it is a *maximum-distance* code that produces the largest possible $D_{\min}$ for any given linear code (N, K) [20]. There are also efficient decoding algorithms that are readily implemented in hardware designs.

7.6.2 Concatenated codes

The block codes described above can be used in conjunction with, or concatenated with, each other, or with the convolutional codes described in the next section. The motivation for using concatenated codes is that the error-correcting performance achieved for a given complexity can be superior to that achieved using a single code alone, largely due to the wider range of different error patterns each code is most suited to correcting in high and low SNRs. When used with interleaving, this approach is also found to have good error-correcting properties for channels in which the errors occur in bursts, a property typical of most fading wireless channels.

The two concatenated codes are usually described as the *inner code* and the *outer code*. Typically a nonbinary code such as an RS code is used as the outer code and a binary block code or convolutional code is used as the inner code. The combined code is then denoted as an (Nn, Kk) code with code rate $R_c = Kk/Nn$ and minimum distance $d = D_{\min} d_{\min}$.

7.6.3 Interleaving

Interleaving is a systematic way of shuffling the order in which the symbols are transmitted over a wireless channel. Usually, the code words are arranged in an interleaving matrix and sent to the modulator in the opposite dimension they were entered. For example, if

the code words were used to populate the rows of the matrix, the symbols would be sent to the modulator in a column-by-column order. For channels in which the errors are not uniformly distributed but instead come in bursts or clusters as might be expected with a sudden deep fade, interleaving results in the corrupted symbols being distributed among multiple code words. Since coding schemes are designed to correct a limited number of symbol errors within a code word, it is reasonable to spread the erroneous symbols over several code words instead of grouping them into a single code word where it is less likely that they can be corrected. Interleaving is a standard part of most modern digital communication systems. It requires no additional overhead bandwidth, but it does introduce a delay since all the code words that populate an interleaving matrix must be received before any can be extracted and sent on to the decoder.

7.6.4 Convolutional codes

The block codes described above are well suited to transmitted information that lends itself to being divided into blocks, such as data packets. However, convolutional codes have actually become more popular, especially for cellular radio applications. This is primarily due to the existence of very simple decoding algorithms (such as the Viterbi algorithm) that can achieve significant coding gain at near optimum theoretical performance. For fixed broadband wireless systems, both block and convolutional codes with be encountered in equipment design.

Convolutional codes can also be described in terms of (n, k) with a code rate $R_c = k/n$. Convolutional codes are distinguished from block codes in that the encoder output is a function not only of a present block of k information symbols but also of the previous information symbols. The dependence on previous information symbols results in the convolutional encoder being a finite state machine (FSM).

A convolutional encoder is shown in Figure 7.18. It consists of L banks of shift registers, each holding k information symbols. At each state of the encoder, the shift

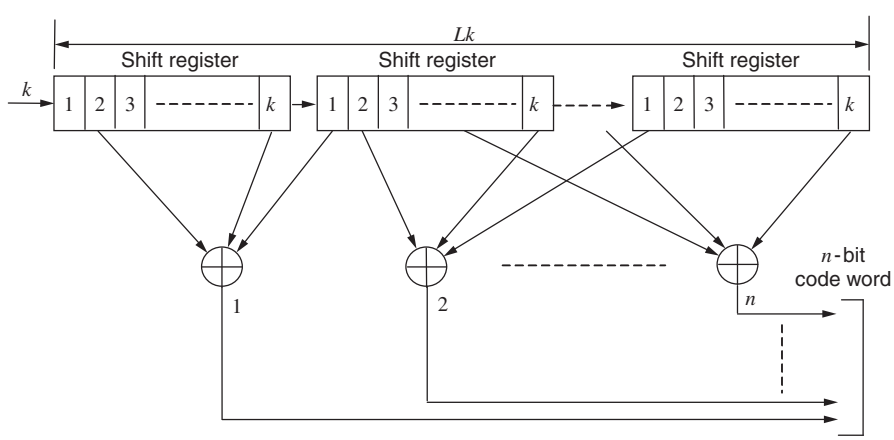

Figure 7.18 A convolution encoder.

registers contain k-bit sequences. The contents of the registers are added together to form the n-bit code word. Which registers are added together to form which code word elements depends on the code generation polynomial.

In addition to the information symbol sequence k and the code word length n, a convolutional code is also described by the constraint length L. The constraint length is simply the number of shift register groups that hold the sequence of information symbols, as shown in Figure 7.18.

Convolutional codes can be decoded in a variety of ways that represent states in a trellis decision tree. The most widely used of these in the Viterbi algorithm [21]. A more detailed discussion of convolutional codes, their encoding and decoding algorithms can be found in several texts such as those in [22,23].

7.6.5 Trellis-coded modulation (TCM)

Trellis-coded modulation (TCM) gained interest in the early 1980s with the published work of Ungerboeck [24–26]. TCM is an amalgamation of coding and modulation that provides a coding gain of 3 to 6 dB without the additional bandwidth necessary for the redundant coding symbols. This was an important innovation that was quickly adopted in many areas of wireless communications.

TCM uses a rate k/n convolutional encoder and directly maps those points onto a set of signal constellation points using a technique called *mapping by set partitioning*. This approach has these basic features:

- A modulation signal constellation (like 64QAM constellation in Figure 7.3b) is used that has more points than would be necessary for the information message alone at the same data rate. The additional signal constellation points essentially increase the modulation efficiency in bits/Hz without increasing the occupied bandwidth.
- The expanded signal constellation is partitioned into subsets in such a way that the Euclidean distance between subsets at each stage of the partitioning is maximized. The Euclidean distance is the magnitude of the multidimensional vector between the signal constellation points in the two subsets.
- The convolutional coding and signal mapping are arranged so that only selected sequences of signal points are used.

As shown earlier in this chapter, modulation with a larger number of signal constellations states will have a larger error rate for a given SNR (E_b/N_0). It seems counterintuitive, then, to use such a constellation. However, this *increase* in error rate is smaller than the *decrease* in error rate that is achieved by using coding among the constellation states. Fundamentally, this is how TCM achieves its advantage.

The performance of TCM has been analyzed in additive white Gaussian noise (AWGN) channels, and channels with flat Rician and Rayleigh fading. From this, rules have been formulated for designing trellis codes for fading channels as set forth in [27].

It should be clear from the above discussion that TCM requires modulation types other than simple binary modulation like BPSK. TCM is well suited to 16QAM, 64QAM, and 256QAM, which are already widely used in high-capacity digital microwave radio. The

use of TCM for these radio systems provides coding gain without sacrificing information throughput to coding overhead.

7.6.6 Coding gain

The error rate improvement achieved through the use of coding is often called the *coding gain*. The coding gain in dB is the reduction of the SNR needed to achieve the same detected/decoded error rate as the uncoded raw error rate. For example, if a SNR of 10 dB is needed to achieve a raw error rate of 10^{-6}, and the same error rate can be achieved with coding and an SNR of 6 dB, the system is said to have a coding gain of 4 dB at an error rate of 10^{-6}. The coding gain will clearly be a function of the code rate and code design, especially the minimum Hamming distance of the code.

If the raw symbol error rate of the channel is denoted as P_M, the code word error rate can be found by viewing the symbols in the code word as independent Bernoulli trials. Independence of adjacent symbols is achieved through interleaving. If the code is guaranteed to correct t symbol errors, then the probability, P_{cw}, a code word will have an error is given by the standard probability for Bernoulli trials [28]

$$P_{cw} \leq \sum_{i=t+1}^{n} \binom{n}{i} P_M^i (1 - P_M)^{n-i} \tag{7.32}$$

where

$$\binom{n}{i} = \frac{n!}{i!(n-i)!} \tag{7.33}$$

Note that the inequality in (7.32) occurs because the code may in some cases be capable of correcting more than t errors but is not guaranteed to do so. The improvement in error rate from coding is clear from (7.32).

When a code word error is made, the symbol error probability is given by [29]

$$P_{MC} = \frac{1}{n} \sum_{i=t+1}^{n} i \binom{n}{i} P_M^i (1 - P_M)^{n-i} \tag{7.34}$$

When the decoded symbols are converted to bits, the bit error probability P_b from (7.34) is

$$P_b = \frac{2^{k-1}}{2^k - 1} P_{MC} \tag{7.35}$$

The net error rate after decoding is often the quantity specified by the system operator and offered to the customers as a specification of the data service. Depending on the service type, this final error rate may be very low, and perhaps only achieved through the use of ARQ even with the improvements available through coding. For real-time services such as voice, ARQ may not be viable in which case the system will rely entirely on the coding for error correction with the expectation that the residual errors will be infrequent enough to go unnoticed by the end user.

With a final net error rate objective, the coding gain can then be used to find the necessary raw error rate, which in turn dictates the minimum SINR. For a fading link, and a given link availability objective, this minimum SINR (and error rate) can only occur a small fraction of the time. The error rate may then be specified as averaged over a 1-min, 5-min, 1-h or similar time frame that is relevant to service users. This leads to the idea of *adaptive coding*, which becomes more robust during fading events. This is similar to the adaptive modulation mentioned earlier in this chapter.

7.6.7 Space-time codes

STCs are codes constructed to exploit the uncorrelated channel response differences that exist when using multiple transmit antenna elements and optionally, multiple receive antenna elements. Multiple-input, multiple-output (MIMO) antenna systems are discussed in Section 6.7.1.

As stated above, coding is the technique of adding redundant symbols to the transmitted signal to improve the ability of the receiver to correctly detect the information data. As described above, conventional coding uses redundancy in the time domain by adding code bits to the time sequence of information bits. A STC uses both the time and space domains to send redundant information to the receiver, the space domain being represented by separate antenna elements distributed in space. This redundancy idea can also be extended to other dimensions of the wireless channel including frequency and polarization, for example.

With STCs, multiple transmit antennas are used to send separate data streams made up of symbols that are organized to achieve performance improvement when detected at the receiver. One of the earliest and simplest examples of an STC is found in the work of Alamouti [30]. This technique will be discussed in detail here for two reasons: (1) it is a straightforward example of link improvement through space-time coding, and (2) the Alamouti method is specified in the draft of the Institute of Electrical and Electronic Engineers (IEEE) 802.16a Standard for 2 to 11 GHz systems as an optional technique for link improvement.

Figure 7.19 shows the Alamouti method using two transmitting antennas and one receiving antenna. This was originally intended as a way to achieve diversity gain at a cellular handset without the burden of adding a second antenna with associated RF/intermediate frequency (IF) circuitry, or additional bandwidth and overhead, to achieve diversity gain. The signals from the two antennas are a sequence of two symbols or symbol blocks, shown as s_1 and s_2 in Figure 7.19. During the first-symbol period at time t, the symbols sent from Antennas 1 and 2, respectively, are s_1 and s_2. During the second-symbol period at time $t + T$, the symbols sent are $-s_2^*$ and s_1^*. The received signals at time periods t and $t + T$ are

$$\begin{aligned} r_1 &= r(t) = h_1 s_1 + h_2 s_2 + n_1 \\ r_2 &= r(t+T) = -h_1 s_2^* + h_2 s_1^* + n_2 \end{aligned} \tag{7.36}$$

where h_1 and h_2 are the channel transfer functions from each of the transmitting antennas to the single receive antenna.

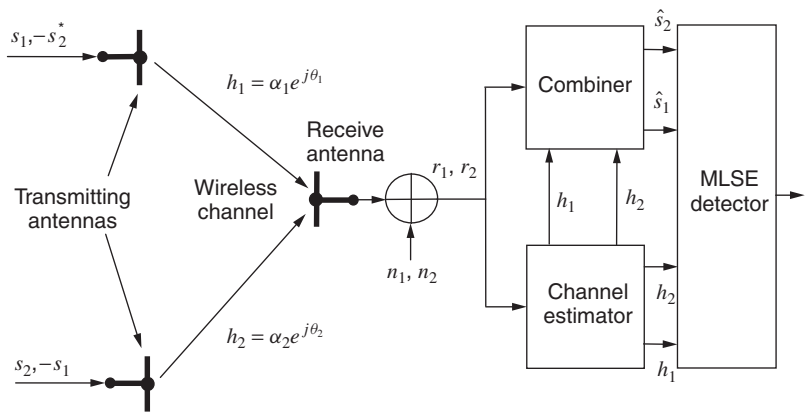

Figure 7.19 Alamouti two-branch transmit diversity with space–time coding.

The combiner produces two estimates of the symbols $\hat{s}_1$ and $\hat{s}_2$

$$\begin{aligned} \hat{s}_1 &= h_1^* r_1 + h_2 r_2^* \\ \hat{s}_2 &= h_2^* r_1 - h_1 r_2^* \end{aligned} \tag{7.37}$$

Inserting the received signals from (7.36) and the channel response from the Figure 7.19 yields

$$\begin{aligned} \hat{s}_1 &= (\alpha_1^2 + \alpha_2^2) s_1 + h_1^* n_1 + h_2 n_2^* \\ \hat{s}_2 &= (\alpha_1^2 + \alpha_2^2) s_2 - h_1 n_2^* - h_2^* n_1 \end{aligned} \tag{7.38}$$

From [30], this is the same result as for a maximal ratio combiner (MRC) with two diversity antennas at the receiver, thus achieving the desired diversity gain result but by using only a single receiving antenna.

In this case the STC is the redundant transmission of the symbols s_1 and s_2 on the two transmitting antennas, but during the second-symbol transmission time, it is reversing the antennas that send the symbols, and sending the complex conjugates of each symbol. The sign of one symbol is also reversed. This STC achieves diversity gain, which in turn will substantially improve the BER in fading channels without additional bandwidth or power. In [30] the concept is extended to multiple receive antennas that provide smaller incremental diversity gain as might be expected. As with any block code, one drawback of the Alamouti scheme is that there is a delay of one symbol time in recovering the two symbols. For a single-channel symbol, this is not a problem, but if the symbols s_1 and s_2 represent long code blocks, this delay could be significant for real-time applications. Also, if the total transmit power remains the same (half to each transmit antenna), the Alamouti scheme also suffers a 3 dB penalty compared to MRC in a receiver using two diversity antennas for reception.

The Alamouti technique is an example of a space-time block code (STBC). The concept of space-time block codes can be generalized to more than two symbols and more than two transmitting antennas. The *space-time code word* for Q length symbol blocks and N transmitting antennas is

$$\mathbf{C}(k) = [\,\mathbf{c}_1(k) \quad \mathbf{c}_2(k) \quad \ldots \quad \mathbf{c}_Q(k)\,] = \begin{bmatrix} c_{11}(k) & c_{12}(k) & \ldots & c_{1Q}(k) \\ c_{21}(k) & c_{22}(k) & \ldots & c_{2Q}(k) \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1}(k) & c_{N2}(k) & \ldots & c_{NQ}(k) \end{bmatrix} \tag{7.39}$$

Each row in (7.39) is a *space-time symbol* and each element a linear combination of possible channel symbols $\mathbf{s}$. Each space-time symbol in $\mathbf{C}$ can be sent in vertical layers as shown in (7.39), or it can be broken into subblocks so that one subblock is sent on each of the antennas. The latter approach is often called *diagonal space-time layering* since the subblocks of each space-time symbol lie on a diagonal across the matrix $\mathbf{C}$.

The received signal $\mathbf{y}(k)$ across a channel with response $\mathbf{H}$ can be written as

$$\mathbf{y}(k) = \sqrt{E_s}\mathbf{H}\mathbf{C}(k) + \mathbf{n}(k) \quad k = 1, 2, 3, \ldots, Q \tag{7.40}$$

where $\sqrt{E_s}$ is the energy per space-time symbol, so that $\sqrt{E_s/N}$ is the energy per constellation point within the code word.

Each symbol block may itself be separately encoded using a block code such as those described in Section 7.6.1. There are various approaches to designing the codes to create the space-time symbols, the most obvious being orthogonal block codes in which the space-time symbols (rows) are orthogonal. The Alamouti two-symbol code uses this approach. The maximum likelihood sequence estimator (MLSE) then decides which symbol was transmitted by calculating the Euclidean distances (analogous to the Hamming distance) between the received code word and the constellation of possible code words and chooses that code word with the smallest distance. From this, error rates can be determined by finding the probability of the incorrect code word being detected, which in turn depends on the Euclidean distance between code words. The derivation of this pairwise error probability is somewhat involved but can be found in [31] and several other publications. For code words $\mathbf{C}$ and $\mathbf{E}$, the probability, averaged over all channel realizations, that $\mathbf{E}$ will be detected when $\mathbf{C}$ is sent is given by an upper bound

$$\Pr(\mathbf{C} \to \mathbf{E}) \le \left(\frac{E_s}{4N_0}\right)^{-r(\mathbf{B}_{C,E})M} \prod_{i=0}^{r(\mathbf{B}_{C,E})-1} \lambda_i^M(\mathbf{B}_{C,E}) \tag{7.41}$$

where M is the number of receive antennas and

$$\mathbf{B}_{C,E} = (\mathbf{C} - \mathbf{E})^T(\mathbf{C} - \mathbf{E})^* \tag{7.42}$$

The term $r(\mathbf{B}_{C,E})$ denotes that the rank of matrix $\mathbf{B}_{C,E}$ and $\lambda_i(\mathbf{B}_{C,E})$ are the nonzero eigenvalues of $\mathbf{B}_{C,E}$. The term N_0 is the noise power so E_s/N_0 is the SNR. The superscript ^{T}

indicates the transpose of the matrix; the superscript * indicates the transpose with complex conjugates of each matrix element.

In addition to block codes, trellis coding can also be used as described in [31]. The design criteria for space-time (ST) trellis codes (STTC) are summarized in [31] as:

- *Rank Criterion*: Maximum diversity is achieved for N transmit and M receive antennas when the matrix $\mathbf{B}_{C,E}$ is full rank for every pair of distinct code words.
- *Determinant Criterion*: With diversity advantage of $N \times M$ as the objective, the minimum of the determinant of $\mathbf{B}_{C,E}$ taken over all distance code word pairs $\mathbf{C}, \mathbf{E}$ should be maximized.

Generally, trellis codes can outperform block codes in terms of error rate and diversity but require increased receiver complexity.

From [30], an original objective of STCs was to provide diversity improvement in narrowband flat-fading channels without adding receive antenna diversity. The best performance improvement is also achieved when the channels are uncorrelated. When these assumptions are not valid; that is, the channel is time-dispersive (frequency-selective rather than flat-fading), and/or correlated to a significant degree, the available performance improvement with STC is more limited. For this reason, ST codes have primarily been of interest when used with OFDM when the cyclic prefix (CP) or zero-padding (ZP) can be inserted to avoid the channel time where multipath signals arrive at the receiver. Both approaches reduce bandwidth efficiency. The percentage reduction can be made small by increasing the block length, but long block lengths are undesirable in a packet-based network with packet acknowledgement retransmission (see Chapter 9).

The ST code performance is described for a single user rather than multiple users in a multiple-access, interference-limited system. In such networks, the link performance will suffer because the number of antennas with interfering signals is increased by N when using STC (even though these antennas are close together and the same total power is distributed among them). As with any modulation/coding scheme, multiuser interference (MUI) can be reduced using beam-forming antennas as described in Section 6.6, but beam-forming requires multiple receive antennas that can only eliminate $M - 1$ interferers [32]. The use of STBC with multicarrier code division multiple access (CDMA) (see Section 8.7.1 for MC-CDMA) where the spreading codes are specifically designed to suppress interference is described in [33].

For fixed broadband NLOS networks, the use of OFDM with STC and MIMO technology is an approach that is being actively pursued. The success of this approach in single link situations has been demonstrated in the field tests (see references for Section 6.7), but at the time of this writing, tests of its performance in high-density MUI environments have not been reported.

7.7 CONCLUSION

There are a wide variety of modulation techniques available to a wireless system designer. Those discussed in this chapter including *M-ary* PSK, *M-ary* QAM, and OFDM are either

the most widely used now or under active development for both LOS and NLOS fixed wireless systems. Several other modulation schemes such as *M-ary* FSK (frequency shift keying) are also available. The choice of a modulation constellation is driven by the required spectral efficiency in bps/Hz as derived from the available channel bandwidth and required channel data rate capacity. For a given link with power and noise limitations, this will also determine the link range for a given availability. Robust systems use adaptive modulation that maintains the link connection at acceptable error rates by reverting to lower level modulation formats and low rate codes when the link is stressed with rain fades, interference, or other impairments.

Errors due to ISI in time-dispersive channels can be effectively combated using both time domain and FDE. There are various trade-offs associated with the use of various equalizers in terms of expected channel delay spread and the severity of the ISI distortion that must be corrected. The appropriate choice of equalizer strategy will therefore depend on the nature of the scattering and multipath, which in turn depends on the nature of the propagation environment.

Coding schemes, especially TCM, provide a means of improving the net system error rate over the raw system error rate by several orders of magnitude. Coding can correct errors resulting from noise as well as interference. (STCs) are one of the more promising approaches that use multiple transmit antennas to exploit the spatial dimension along with the time dimension of the wireless channel to achieve diversity gain and improved error rate performance.

Modern fixed broadband wireless systems use many of these techniques in various combinations to achieve high reliability links. Understanding the parameters involved in modulation, equalizers and coding, and how they impact the system performance, is important for successful wireless system design.

7.8 REFERENCES

[1] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill. 2nd Edition. 1989.
[2] J.G. Proakis and M. Salehi. *Communications System Engineering*. Upper Saddle River: Prentice-Hall. 1994. page 553.
[3] G.L. Stuber. *Principles of Mobile Communication*. Boston: Kluwer Academic Publishers. 1996.
[4] R. Van Nee and R. Prasad. *OFDM for Wireless Multimedia Communications*. Boston: Artech House Publishers. 2000.
[5] H. Sari, G. Karam, and I. Jeanclaude, "Frequency domain equalization of mobile radio and terrestrial broadcast channels," *IEEE Global Telecommunications Conference*, San Francisco, pp. 1–5, November, 1994.
[6] H. Sari, G. Karam, and I. Jeanclaude, "An analysis of orthogonal frequency-division multiplexing for mobile radio applications," *Proceedings of the 44th IEEE Vehicular Technology Conference*, Stockholm, pp. 1635–1639, June, 1994.
[7] R. van Nee and R. Prasad. *OFDM for Wireless Multimedia Communications*. Boston: Artech House Publishers. 2000. pp. 120 ff.
[8] T. Wilkinson and A. Jones, "Minimisation of the peak to mean envelope power ratio of multicarrier transmission schemes by block coding," *Proceedings of the 45th IEEE Vehicular Technology Conference*, Chicago, pp. 825–829, July, 1995.

[9] R. Van Nee. "OFDM codes for peal-to-average power reduction and error correction," *IEEE Global Telecommunications Conference*, London, pp. 740–744, November, 1996.
[10] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill. 2nd Edition. 1989. p. 262.
[11] R.F. Pawula, S.O. Rice, and J.H. Roberts. "Distribution of the phase angle between two vectors perturbed by gaussian noise," *IEEE Transactions on Communications*, vol. COM-30, No. 8, pp. 1828–1841, August 1982.
[12] K. Murota and K. Hirade, "GMSK modulation for digital mobile radio telephony," *IEEE Transactions on Communications*, vol. COM-29, No. 7, pp. 1044–1050, July 1981.
[13] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill. 2nd Edition. 1989. page 284.
[14] J.G. Proakis and J. Miller, "An adaptive receiver for digital signaling through channels with intersymbol interference," *IEEE Transactions on Information Theory*, vol. 15, pp. 484–497, July 1969.
[15] M. Schwartz, W.R. Bennett, and S. Stein. *Communication Systems and Techniques*. New York: McGraw-Hill, 1966. page 349.
[16] R.W. Lucky, "Automatic equalization for digital communications," *Bell System Technical Journal*, vol. 44, pp. 547–588, April 1965.
[17] G.L. Stuber. *Principles of Mobile Communication*. Boston: Kluwer Academic Publishers. 1996. page 265.
[18] D. Falconer, S.L. Ariyavisitakul, A. Benyayamin-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.
[19] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill. 2nd Edition. 1989. p. 386 ff.
[20] R. Steele. *Mobile Radio Communications*. London: Pentech Press. 1992. page 424.
[21] A.J. Viterbi and J.K. Omura. *Principles of Digital Communications and Coding*. New York: 1979.
[22] G.L. Stuber. *Principles of Mobile Communication*. Boston: Kluwer Academic Publishers. 1996.
[23] J.G. Proakis and M. Salehi. *Communications System Engineering*. Upper Saddle River: Prentice-Hall. 1994.
[24] G. Ungerboeck, "Channel coding with multilevel phase signals," *IEEE Transactions on Information Theory*, vol. 28, pp. 55–67, January, 1982.
[25] G. Ungerboeck, "Trellis coded modulation with redundant signal sets- Part I: Introduction," *IEEE Communications Magazine*, vol. 25, pp. 5–11, February, 1987.
[26] G. Ungerboeck, "Trellis coded modulation with redundant signal sets- Part II: State of the art," *IEEE Communications Magazine*, vol. 25, pp. 12–21, February, 1987.
[27] G.L. Stuber. *Principles of Mobile Communication*. Boston: Kluwer Academic Publishers. 1996. page 367.
[28] P.Z. Peebles. *Probability, Random Variables, and Random Signal Principles*. New York: McGraw-Hill. 1980. page 72.
[29] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill. 2nd Edition. 1989. page 430.
[30] S.M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October, 1998.
[31] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communications: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March, 1998.
[32] A.F. Naguib, N. Seshadri, and A.R. Calderbank, "Applications of space-time codes and interference suppression for high capacity and high data rate wireless systems," *Proceedings of the 32nd Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, pp. 1803–1810, November, 1998.
[33] Z. Lui and G.B. Giannakis, "Space-time block-coded multiple access through frequency-selective fading channels," *IEEE Transactions on Communications*, vol. 49, no. 6, pp. 1033–1044, June, 2001.

8

Multiple-access techniques

8.1 INTRODUCTION

The wireless communication spectrum from approximately 150 kHz to 100 GHz is segmented into bands that have been designated for use by particular kinds of wireless services. These designations have been established by the International Telecommunications Union (ITU) on a worldwide basis, with further spectrum partitioning instituted by countries within these broader spectrum designations. Dividing the spectrum into blocks so that a multitude of users can have simultaneous access to it without conflict is the most basic form of frequency division multiple access (FDMA). While global spectrum partitioning is the mechanism for multiple access to the entire spectrum, for wireless systems operating in a particular geographic area in the same frequency band, multiple access means the simultaneous access to the same frequencies or channels within the same area by multiple users. In this context, 'simultaneous' means the system end user perceives the service to be continuous and immediate whether or not the wireless system is actually providing information to that user on a continuous and immediate basis. Service that is not perceived as continuous or immediate by the user (blocked or dropped calls in a cellular system, for example), or is otherwise flawed, is considered to have low quality of service (QoS). The number of simultaneous users that can be supported at a given QoS level is one measure of the capacity of the system and therefore directly tied to its success in achieving its service and commercial objectives. The capabilities of the system's multiple-access technology is therefore of fundamental importance in determining whether these objectives can be met.

While the multiple-access QoS delivered to the user is a function of the service type and acceptable service delay, the radio channel performance criteria for all multiple-access schemes can be reduced to the fundamental consideration of interference between systems (intersystem) and within systems (intrasystem). Controlling intersystem interference

Fixed Broadband Wireless System Design Harry R. Anderson
© 2003 John Wiley & Sons, Ltd ISBN: 0-470-84438-8

is usually the role of administrative rules or industry standards since the system operators competing for use of the spectrum will not necessarily conduct themselves in a cooperative way. The systems competing for use of the spectrum could be operating in adjacent geographic areas or could be operating in the same geographic area or market.

By contrast, intrasystem interference is caused by transmitting nodes within the network over which the operator has complete control. Within the broader context of avoiding intersystem interference and complying with any mandated modulation types, channelization formats, and so on, the system operator has considerable flexibility in choosing technology and multiple-access techniques to maximize system capacity and service levels. Since the greatest advantage can be gained for the operator through an intelligent system design to enhance capacity, intrasystem multiple access will be the primary subject of this chapter.

As noted in [1], the radio spectrum is a natural resource that is distinctly different from other natural resources such as forests, petroleum reserves, water sources, and so on. It is not depleted through use – extraordinarily, it is only wasted when it is not being used. The most efficient spectrum usage is achieved by maximizing the occupancy of time, frequency, and space within a designated frequency band.

8.1.1 Intersystem multiple access

Within a given frequency band in the same general geographical area, there may be several users who seek to operate their systems without destructive interference from other users in the same frequency band and area. The criteria for when destructive interference occurs are often written into the administrative rules for the spectrum such as those found in the Rules of the Federal Communications Commission (FCC). These rules can be quite detailed, such as those governing two-way multipoint, multi-channel distribution service (MMDS) systems in the United States where the specific propagation model, databases, system description, calculation methods, and interference thresholds are specified in detail. These specific methods must be used in order to be granted authorization to build a network in the MMDS spectrum (2.5–2.7 GHz) [2]. In contrast, the FCC Rules for intersystem interference for cellular and personal communications service (PCS) systems only describe power density limits beyond the boundaries of the system service limits.

Besides government statutes, there are also industry standards that have been established to accomplish harmonious sharing of spectrum among various users. Again, these can be quite detailed like the Telecommunication Industry Association (TIA) TSB-88-A for calculating interference and spectrum compatibility between mobile radio systems [3]. Other industry standards such as the interoperability standard for 10 to 66 GHz developed by the Institute of Electrical and Electronic Engineers (IEEE) 802.16 Working Group do not specify parameters for interference calculations but simply leave much of this to rely on ‘good engineering practice’ that can certainly lead to significant differences of opinion about how interference should be calculated and whether it is ‘destructive’ [4].

The spectrum administration rules and industry standards are all nonhardware mechanisms to achieve successful FDMA among various spectrum users. They are driven by the capabilities of the technology to the extent that the rules can allow tighter spectrum sharing when the technology employed by the operators is less susceptible to external interference. For example, at various times in the past the FCC has mandated the use of equipment

and antennas that meet certain technical performance thresholds (out-of-band emissions, modulation efficiency, receiver bandwidth, antenna directionality, etc.) to increase the number of users and systems that can be accommodated within a spectrum band. Spectrum capacity and how technologies affect capacity is discussed further in Chapter 12.

8.1.2 Intrasystem multiple access

A given fixed broadband network may have one or more users all of whom require service of a given capacity, quality, and consistency. The capacity, quality, and consistency may be embodied in a contractual commitment between the service operator and the end user. The system capacity and the extent to which the system can accommodate the predicted traffic levels (and requisite service quality) are the primary design drivers. The way in which the available spectrum is partitioned for sharing by multiple users, and the multiple-access technologies that are then used, directly affects the traffic volume that can be supported. The capacity and quality issues ultimately determine whether the system can meet the commercial objectives of the system operator who seeks a system that is profitable, sustainable, and scalable to serve increasingly dense and dispersed demand.

In Chapters 7 and 12, the concept of adaptive modulation is discussed in which the modulation type is adjusted to respond to the conditions of the wireless channel. Multiple access can also be adapted in the same way to fulfill the requirements of service agreements with the end users. Wireless operators can construct several service levels with different price structures that offer different levels of access. These service levels vary from top-level service in which the end user is provided with immediate best access to the network when requested to lower-level access that is often described as 'best effort'. The Internet is an example of a network with a 'best effort' service objective. An adaptable multiple-access scheme is therefore not just a technical mechanism to increase capacity but is also a fundamental part of the system operator's business plan.

8.1.3 Duplexing

Duplexing is the term given to the method used to accomplish two-way communication between terminals of a point-to-point link or between a network hub and a remote terminal. Two methods to duplexing are used in fixed broadband wireless – frequency division duplexing (FDD) and time division duplexing (TDD). Unless mandated by administrative rules (which is the case for many services), the choice of the duplexing method is generally independent of the multiple access and modulation techniques used.

FDD is currently more widely used for fixed broadband wireless. With FDD a separate frequency is assigned for hub-to-remote communication (downstream or downlink) and for remote-to-hub (upstream or uplink) communication. The frequency separation between uplink and downlink channels is usually chosen to be great enough so that construction of suitable filters to separate the two signals at a given terminal is practical. FDD is the approach used with all 1G, 2G, and most 3G cellular systems, as well as most other mobile communications. However, when the volume of downlink and uplink traffic is not symmetric as may be the case with nonvoice data services, FDD results in an underutilized spectrum, usually on the uplink, which is inefficient. This has lead to an increased interest in TDD in recent years.

TDD uses the same frequency or circuit for both uplink and downlink communications between the network hub and remote terminal. In the broader wireless communications world, TDD is referred to as *simplex* transmission; the 'TDD' appellation has come into vogue only recently. TDD is actually the oldest form of duplexing, having been used for telegraph communications some 150 years ago, in which an operator would transmit a message and then listen for a response. Simplex or TDD is also used in many current-day communication systems such as aircraft communications in which all aircraft and the airport tower use the same frequency to send and receive voice messages. TDD potentially can make more constant and efficient use of a radio channel, especially for asymmetric uplink–downlink traffic flow. However, because both hub stations and remote transmitters can potentially use the same frequency, the interference considerations, and synchronization of transmissions to avoid interference, become more complex. This can impact the extent to which frequencies may be reused within a system.

The technical issues associated with FDD and TDD will be discussed in this chapter with particular consideration of spectrum efficiency and how interference calculations should be carried out.

It should also be mentioned that duplexing can be accomplished using multiple delivery technologies. For example, some proposed broadband systems use relatively high-speed satellite transmission or television broadcast channel transmission for the downlink and a relatively slow-speed telephone line for the uplink, again motivated by the assumption that the downlink traffic volume is greater than the uplink traffic volume. Because such hybrid systems are not entirely wireless, they will not be considered further in this book.

8.2 FREQUENCY DIVISION MULTIPLE ACCESS (FDMA)

With FDMA, the spectrum available to the system operator is segmented into frequency channels. A frequency channel or slot is a 'physical' channel that is identified by its center frequency and bandwidth. When a remote terminal is installed, or initiates communication, a frequency channel is assigned to support the communication between the system hub and that terminal. If the system uses FDD, both an uplink and downlink channel are assigned. If the system uses TDD, only a single channel is assigned for both uplink and downlink. The bandwidth of the frequency slots can be uniform or variable. In many cases, administrative regulations, and sometimes industry standards, mandate the channelization scheme so that all channels are of uniform width. Figure 8.l shows a simple diagram of an FDMA with N channels that are occupied full time. Figure 8.2 shows the traditional view of the spectrum axis with the radio channels of varying bandwidths and a representation of the power spectral density (PSD) within each allocated channel.

The channel assignment process can be done statically as part of the planning process in which a given frequency is assigned to a terminal when the terminal is installed. The channel may also be assigned dynamically whenever the terminal initiates communication. This is the standard approach used in cellular systems. The assignment criteria for selecting a channel include interference avoidance and, in systems with bandwidth

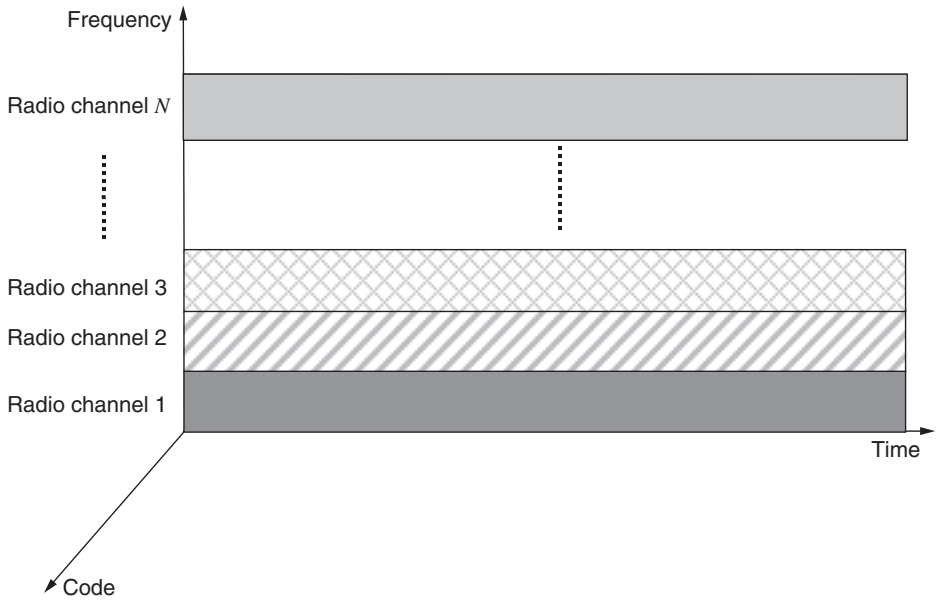

Figure 8.1 Frequency division multiple access (FDMA).

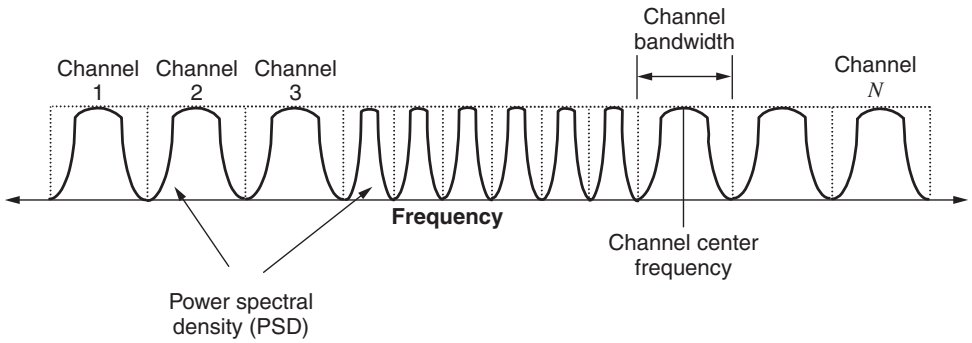

Figure 8.2 FDMA channelization of allocated spectrum with different occupied bandwidths.

flexibility, adequate bandwidth to support the data rates required by the terminal. An effective channel assignment process is fundamental for achieving efficient use of the spectrum. Accordingly, frequency assignment and planning will be discussed in some detail in Chapter 12 for particular network configurations.

Essentially all wireless systems use FDMA in some form, even systems that are normally referred to as *time division multiple access* (TDMA) systems. For example, the IS-136 TDMA cellular system used in the United States uses the same FDMA channelization scheme as the first generation advanced mobile phone service (AMPS) cellular

system; it additionally allows multiple users per channel by dividing the channel into assigned time slots (TS). A similar approach is used for the global system for mobile communications (GSM) cellular system deployed in many parts of the world: 124 frequency channels are shared by dividing each into 8 assigned TS. Therefore, FDMA should be recognized as a 'first level' multiple-access approach used to some extent by all wireless technologies with further multiple access achieved through 'second level' spectrum partitioning of each FDMA channel using TDMA, code division multiple access (CDMA), or space division multiple access (SDMA). TDMA, CDMA, SDMA, and variations of these techniques are discussed in the following sections of this chapter.

8.2.1 FDMA interference calculations

The interference in an FDMA channel can come from three sources:

- Cochannel interference from other transmitters occupying some or all of the frequencies of the receiving channel currently used by the victim receiver.
- Adjacent channel interference from transmitters occupying some or all of the frequencies either above or below the receiving channel of the victim receiver.
- Spurious interference from intermodulation products and other, often unidentifiable, sources.

Spurious interference can usually be controlled to an acceptable level by proper installation techniques and equipment usage and is thus not a part of the design process. This section will focus on cochannel and adjacent interference.

The signal-to-interference + noise ratio or SINR, is the fundamental quantity that usually serves as a design objective for determining whether the interference level is acceptable. SINR is given by

$$\text{SINR} = \gamma = \frac{S}{I + N} \tag{8.1}$$

Each of the variables in (8.1) is a random process. Calculation of the noise power, N, and the interference power, I, in (8.1) is discussed in the following sections.

8.2.1.1 Noise power

The noise voltage N is usually assumed to be a zero-mean Gaussian random variable, with the power given by the variance of its amplitude distribution. The noise can come from a variety of sources but for frequencies above 2 GHz, the dominant noise source is thermal noise in the receiver circuitry and antenna system. The noise contribution from these sources is characterized by the *noise figure*. A perfect device that contributes no noise to the system has a noise figure of 1 or 0 dB. For other values of noise figure, the noise power density η (variance of the Gaussian distribution) is given by

$$\eta = (F - 1)K_b T_0 \quad \text{(Watts/Hz)} \tag{8.2}$$

where F is the noise figure, K_b is Boltzmann's constant $= 1.37 \times 10^{-23}$, and T_0 is the ambient temperature in degrees Kelvin of the environment in which the device operates; it is usually taken to be 290°. The noise of the device can also be characterized by its effective *noise temperature* T_e, which is related to noise figure by

$$T_e = (F - 1)T_0 \tag{8.3}$$

For very low noise devices such as cryogenically cooled amplifiers, the effective noise temperature is often the preferred way to characterize the noise contributed by the device.

The noise power density given by (8.2) can be used to find the total noise power N at the input of the receiver in a given bandwidth by simply multiplying by that bandwidth B

$$N = (F - 1)K_b T_0 B \tag{8.4}$$

The bandwidth used here is called the *equivalent noise bandwidth*, a uniform rectangular bandwidth that produces the same noise power as the actual receiver filter bandwidth. The noise figure in (8.4) is the noise figure for the entire receiving system, although the first stage of a receiver is usually the most significant contributor to the total system noise. Using the manufacturer's noise figure for the receiving equipment for the relevant bandwidth provides the system *noise floor* via (8.4). A 6-MHz MMDS system with a 4-dB noise figure will have a noise floor of −134.4 dBW or −104.4 dBmW. Link thermal fade margins are calculated by comparing the mean received signal power with the noise floor power.

The signal, interference, and noise powers in a receiver system, and the resulting SINR value, are normally referred to the input terminals or connector of the receiver. The transmission line loss is taken into account by appropriate attenuation of the signal and interference powers. Alternately, the SINR calculation could be referred to the antenna terminals in which case the transmission line loss adds directly to the noise figure and calculated noise power, but the signal and interference are not attenuated by the transmission line loss. The two approaches result in the same value of SINR.

Since the transmission line loss directly decreases the signal-to-noise ratio (SNR) [but not signal-to-interference ratio (SIR)], it is important to reduce its length as much as practically possible for traditional wireless systems in which tower-mounted antennas are some distance from the transmitter or receiver equipment. As described in Chapter 6, recent equipment developed for LMDS and other fixed broadband systems designed for building rooftop installation have outdoor units (ODUs) in which the transmitter power amplifier or receiver preamplifier modules are directly connected to the antenna, thus eliminating the radio frequency (RF) transmission line or waveguide entirely.

The noise power is just one moment (the variance) of the noise amplitude distribution. As shown in Section 7.4.1, errors in signal detection occur when the noise amplitude exceeds a relative signal level depending on the modulation type. It is the noise amplitude distribution, then, not just the variance that determines when errors occur. Thermal noise amplitude is modeled as a Gaussian distribution; atmospheric and manmade noise have different, more impulsive, amplitude distributions that result in different error values for the same average noise power.

8.2.1.2 Cochannel and adjacent channel interference

The interference term I in (8.1) is a random process that is the sum of the cochannel and adjacent channel interference

$$I = \sum_{n=1}^{N} I_{co,n} + \sum_{m=1}^{M} I_{adj,m} \tag{8.5}$$

where N is the number of cochannel interferers and M is the number adjacent channel interferers. The amplitude of the interferers I_n and I_m may be relatively constant most of the time as with LMDS systems using narrowbeam antennas, or they may be fading signals that are best described as compound Rician or Rayleigh fading processes with lognormal-distributed mean values (see Chapter 4). Such descriptions for interference are applicable to non-line-of-sight (NLOS) fixed broadband systems.

As mentioned in the introduction to this chapter, the interfering signals can occupy the same bandwidth slot as the desired signal S or a bandwidth slot that overlaps the desired signal bandwidth slot to some extent. This situation is illustrated in Figure 8.3 in which the receiver filter shape and interfering PSD are depicted. The amount of interference intercepted by any arbitrary receiver bandwidth from any arbitrary interfering PSD can be found by integrating the overlap area. This approach is sufficiently general that it can be used for either cochannel or adjacent channel interference; indeed these descriptions of interference become somewhat meaningless when considering the wider variety of possible desired receiver bandwidths, center frequencies, and interfering PSDs than may be used within a given network.

The mean power from each interferer is first calculated at the terminal of the victim receiver taking into account the interferer's radiated power in the direction of the victim receiver, the total mean propagation path loss, and the victim receiver antenna gain in the direction of the interferer. That power is then further reduced by the extent to which the receiver filter reduces the intercepted power based on the interferer's PSD. The net mean interference power P_i resulting from this process, is given by

$$P_i = P_T + G_T(\varphi_T, \theta_T) + G_R(\varphi_R, \theta_R) - L_p - A_F - A_S \text{ dBW} \tag{8.6}$$

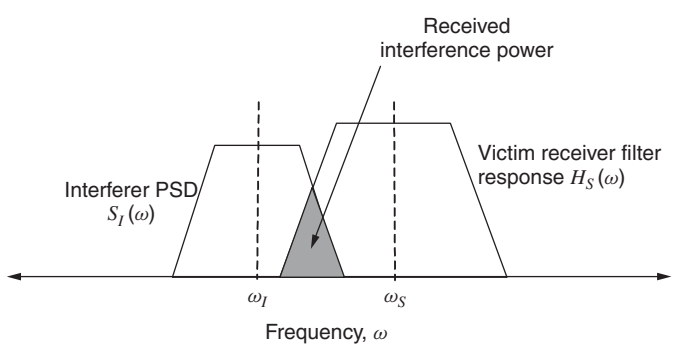

Figure 8.3 Received interference for arbitrary desired and interfering frequencies and occupied bandwidths.

where

P_T = transmitter power in decibel-watts.

$G_T(\varphi_T, \theta_T)$ = the gain in decibels of the transmitting antenna in the direction of the victim receiver where φ_T is the horizontal plane angle (azimuth) and θ_T is the vertical plane angle (elevation).

$G_R(\varphi_R, \theta_R)$ = the gain in decibels of the victim receiving antenna in the direction of the interferer where φ_R is the horizontal plane angle (azimuth) and θ_R is the vertical plane angle (elevation).

L_p = mean path loss in decibels, which includes free space loss, clutter loss, reflections, diffraction, and multipath.

A_F = the attenuation factor in decibels due to the receiver filter rejecting some or all of the interfering power.

A_S = the attenuation in decibels due to other system losses such as transmission line or waveguide losses, radome losses, and so on

The transmitting power level, antenna gains, and mean path loss will be specific to each interferer. The system losses A_S will be common to all received signals including the desired signal. From Figure 8.3, the receiver filter attenuation A_F is given by

$$A_F = 10 \log \left[\frac{\int_{-\infty}^{+\infty} H_S(\omega - \omega_s) S_I(\omega - \omega_I)\, d\omega}{\int_{-\infty}^{+\infty} S_I(\omega)\, d\omega} \right] \tag{8.7}$$

where $H_S(\omega)$ is the desired signal receiver filter response center on frequency ω_s and $S_I(\omega)$ is the PSD of the interfering signal centered on frequency ω_I. As a practical matter, (8.7) is usually performed as a numerical summation using receiver filter response and transmitter PSD curves provided by the equipment manufacturer. Examples of some actual transmitter PSD and receiver filter response curves are shown in Figures 8.4 and 8.5, respectively.

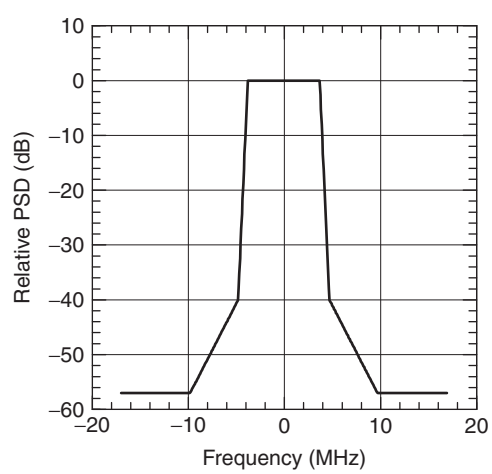

Figure 8.4 Example of a transmitter power spectral density (PSD) envelope.

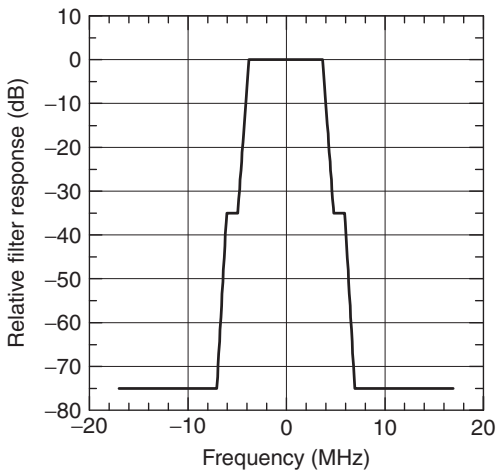

Figure 8.5 Example of a receive filter response envelope.

These curves are the envelopes of PSD and receive filter responses for LMDS equipment that are developed from the measured responses of sample equipment taken from the production process.

Since the receive filters and transmitter power spectra are not ideally band-limited, FDMA systems require a *guard band*, which is usually provided for by making the allocated channel bandwidth wider than that necessary to accommodate the intended signal spectrum width. *Spectrum masks*, similar to the PSD curve in Figure 8.4, define the required transmitter power limits as a function of frequency. If the administrative or industry standard spectrum masks are defined, they can also be used for interference calculations instead of measured equipment performance data.

8.2.1.3 Multiple interferers in LOS networks

As discussed in Chapter 7, the amplitude of the interference itself can actually be subject to fading due to changing propagation path conditions, or other time variations such as power increases resulting from transmitter power control action when the interfering link itself is being subjected to some fading event like rain. The statistics for the variations that might be expected are generally not known. The attempts to model variations in interfering signal amplitudes as Rayleigh or Rician are approximations based on observations in mobile radio systems that can be significantly wrong for line-of-sight (LOS) fixed broadband systems. Moreover, as indicated in (8.5), the total interference can actually be the sum of several interferers, each with different mean values from (8.6) and different statistics describing the amplitude variations.

Owing to the lack of specific information about the fading properties of the interferers, the standard approach is to simply consider the interferers as additional system noise and by doing so, approximate their envelopes as Rayleigh-distributed like Gaussian noise. For

M interferers of arbitrary center frequency and PSD, the denominator in (8.1) is then given by

$$I + N = \sum_{m=1}^{M} P_m + P_N \tag{8.8}$$

where P_m is given by (8.6) for each interferer. For power changes that may be predicted, such as power increases due to interferer transmitter automatic power control (APC) action, the mean interference power is increased accordingly and the SINR calculation in (8.1) is performed under these conditions. Of course, the desired system transmitter APC may also be capable of responding to the increased interference. *Link coupling* interactions of this type will be considered in Chapter 11.

The quality of the approximation in (8.8) where interferers are considered as additional noise will be increasingly poor as the interfering power is less and less 'noiselike'. For multiple fading interferers of comparable amplitude, modeling the combined interference as additional noise power is a reasonable approach. The worst case for this approximation occurs when the total interference consists of a single nonfading interferer. Even if the multiple interferers are present, if a single interferer dominates the others by several decibels, the envelope amplitude of the combination will be more Rician than Rayleigh.

The worst-case quality of the approximation in (8.8) can be examined using the analysis presented in Chapter 7 for the error rate performance of quadrature phase shift keying (QPSK) and 16QAM signals being detected in the presence of a single sine wave interferer plus Gaussian noise. Consider two cases for QPSK: one has noise and interference powers of equal amplitude and is 10 dB weaker than the desired signal and the second has the interfering signal removed and the noise power increased to 7 dB weaker than the desired signal. The second case is equivalent to considering the single interferer as additional noise as given in (8.8). The error probability P_e in the first case from Figure 7.13 is about 2.5×10^{-3} while in the second case the error probability from Figure 7.7 is about 1.6×10^{-3}. The results from a similar comparison for 16QAM at SNR = 15 dB and SIR = 15 dB level using Figures 7.10 and 7.8 are included in Table 8.1.

From this comparison, using the noise approximation for interference provides P_e values that are close to or higher than the correct P_e values for the single sine wave + noise case. Similar comparisons done numerically using the equations in Chapter 7 show that the 'all noise' approximation provides a worst-case estimate of P_e for SNR and SIR values of practical interest, that is, values that would be applied to realistic system designs. Therefore, using the noise approximation in (8.8) for combined interference constitutes a conservative approach to system design.

Table 8.1 P_e for the noise approximation of interference for C-QPSK and 16QAM

Modulation	SNR,SIR level	Single sine wave + noise	All noise approximation
C-QPSK	10 dB	2.5×10^{-3}	1.6×10^{-3}
16QAM	15 dB	1.0×10^{-4}	7.0×10^{-4}

8.2.2 Spectrum utilization

Pure FDMA systems have reserved channels for the various communication links in the network. This is efficient spectrum utilization only if the service on the link requires full-time access, as may be the case with 'real-time' services such as video or audio transmissions or building-to-building business data links. However, even these signals can be stored and delayed so that time-sharing the channel with another service is still possible. From the perspective of the end user, the service still appears to be uninterrupted. The memory capabilities of the transmitting and the receiving equipment accessing the channel therefore become an important factor in determining the flexibility in choosing a multiple-access scheme. The delay that can be tolerated, and the somewhat nebulous definition of 'real time', also play a role. Clearly, an entire sporting event could be stored and transmitted incrementally on a shared channel so that it appears as a continuous real-time transmission, but the event, or critical parts of the event, may have finished long ago, thus destroying the immediacy of what the end user experiences.

The issues associated with service type and when immediate transmission versus deferred transmission is acceptable are discussed in more detail in Chapter 9. For the present purpose, it is sufficient to say that with current technology, it is a very rare service that truly requires full-time continuous access to a communications channel. In recognition of this fact, alternative schemes such as TDMA and CDMA have been developed, which can achieve greater spectrum utilization for most services through shared multiple access to the frequency band without deteriorating the apparent real-time immediacy of the end user experience.

8.3 TIME DIVISION MULTIPLE ACCESS (TDMA)

Within a specific radio channel, multiple users can be provided with service if the times during which transmissions are sent and received by each user are controlled and coordinated by the system so that interference conflicts do not occur. With TDMA systems, access to the radio channel is typically divided into TS that are then assigned to individual users, much as full-time radio channels would be assigned to users in a purely FDMA system. The TDMA approach is illustrated in Figure 8.6. Each radio channel is divided into TS that are synchronized within a radio channel on a given base station or hub, but are not necessarily synchronized from channel to channel or from hub to hub on a given channel.

As a second-level multiple-access scheme, TDMA provides another degree of freedom that allow greater flexibility for accommodating service types. For example, a user that requests a service that has high data rate demands could be assigned multiple TS on a given channel, or across multiple channels, to increase the total data throughput for information transmitted to that user, up to full usage of all available TS with the concomitant denial of service to other users. Flexing the number of TS assigned to a user on the basis of requested service is the approach employed in General Packet Radio System (GPRS) and Enhanced Data rates for GSM Evolution (EDGE) cellular radio systems. It is another form of link adaptation mentioned in Chapter 7.

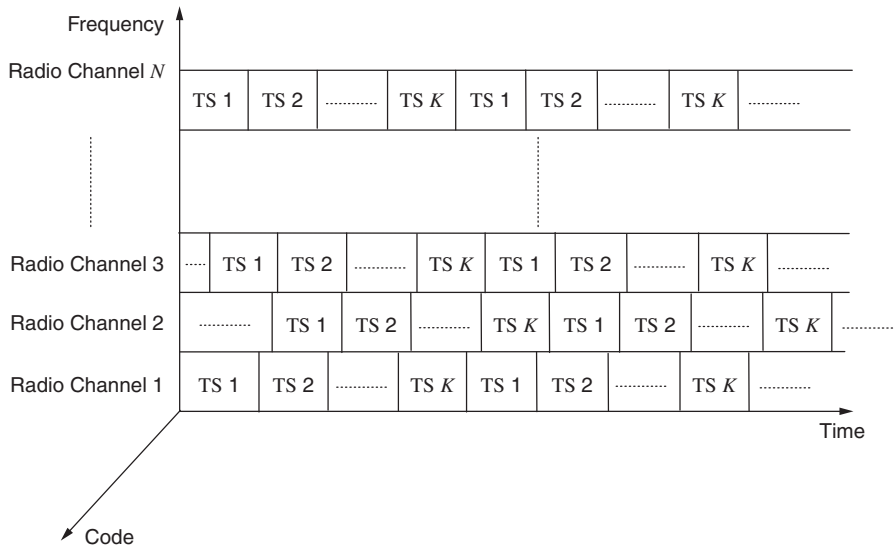

Figure 8.6 Time division multiple access (TDMA) with schematic of time slots (TS).

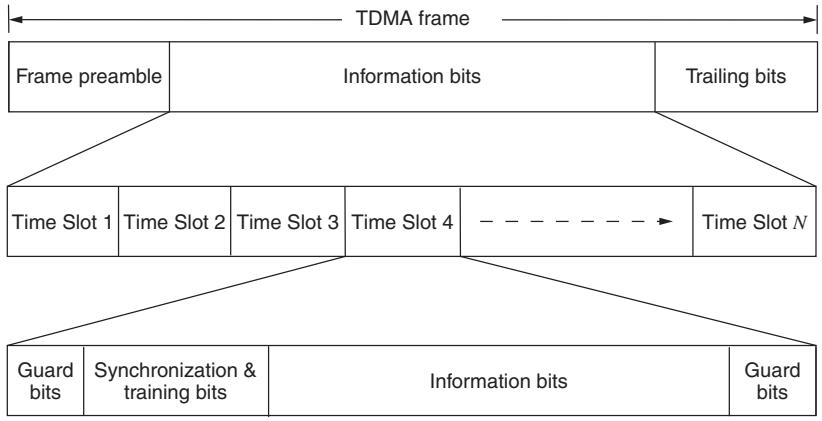

Figure 8.7 Generic TDMA frame and time slot structure.

The TS shown schematically in Figure 8.6 are actually grouped together in *frames* to ensure that they are coordinated with one another. Figure 8.7 shows a generic TDMA frame with preamble and trailing bits. The information frame is divided into TS that are self-contained *bursts* of data including training and synchronization bits. The training bit sequence is known to the terminal and used to train the equalizer as discussed in Section 7.4.4. Depending on the modulation type, these bits may also be used to establish a synchronous reference to facilitate data symbol detection.

Essentially any modulation type may be used with TDMA to provide the required data rate within the constraints of the link budget and design SINR. The GSM cellular system is a TDMA system that uses Gaussian minimum shift keying (GMSK) modulation, while EDGE, the higher data rate 2.5G version of GSM, uses 8PSK for a higher bits/symbol ratio. 16QAM or 64QAM are used in some forms of TDMA for fixed broadband systems.

The generic TDMA frame structure shown in Figure 8.7 may vary from system to system depending on what is to be accomplished, but all TDMA systems must dedicate a certain portion of the transmitted symbol stream to synchronization and timing and guard bits or symbols to prevent overlap of consecutive frames or consecutive TS. The overhead associated with these functions is in addition to the overhead for coding described in Section 7.5. The TDMA overhead bits will be system-specific, but can range from 10 to 30% of the total data throughput capacity. Therefore, overhead in TDMA systems is an important consideration in assessing the information payload that can be delivered.

The term 'channel' is most commonly applied to a radio channel with a particular center frequency and occupied bandwidth as with an FDMA system. However, a sequence of TS utilized by a terminal is also referred to as a *channel* in the rudimentary sense of the physical mechanism that conveys the communicated information. In system protocol nomenclature, the frequency and time physical channels are referred to as the physical channel layer or *PHY* layer.

Frequency and time physical channels are also filled with information that make up different 'logical' channels. Logical channels make up part of what is called the *medium-access control* or *MAC* layer. The MAC layer logical channels and protocols are properties of the system standard. While the MAC layer is important for messaging and services support, fixed broadband wireless system design and planning primarily involves the successful operation of the radio or PHY layer. For this reason, except in a few specific areas, most aspects of the MAC layer design will not be dealt with in this book.

8.3.1 TDMA intercell interference

From Figure 8.6 it is clear that within each radio channel the assignment and use of the TS by each terminal is coordinated so that the signals arriving from each terminal fall in the correct TS. This requires the hub to send a signal to each terminal on a given channel that assigns a *timing delay* or advance to that terminal, so the burst transmission from each arrives at the correct time. The required delay will be largely a function of the distance from the hub to the terminal, which can be deduced from the round-trip transmission delay. Multipath signals may introduce some uncertainty in this process. So that the signal from each terminal occupies its assigned time slot, the timing adjustments τ_k are used as shown in Figure 8.8 for a multicell system.

The time delay adjustments are made at each hub to prevent overlap of TS used by the terminals with which it is communicating; however, this synchronization does not extend to the timing of TS used on the same radio channel at nearby cells. The downlink interference from neighboring hubs at a remote terminal can be readily calculated since the hub position and path loss to a terminal location are known, and it is reasonable to assume that all TS on a radio channel are going to be occupied. For the uplink interference, however, the interfering signals on a radio channel are originating from remote terminals

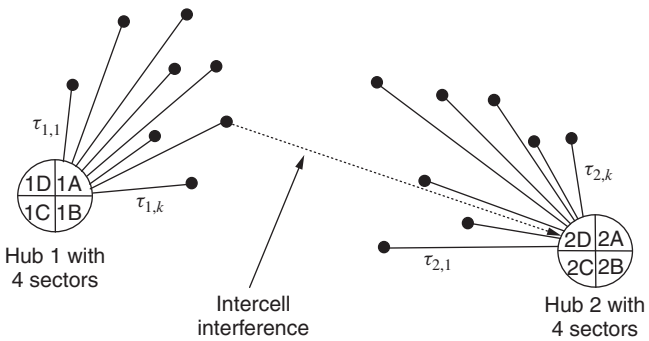

Figure 8.8 Remote terminals in a TDMA system with time delays τ_k adjusted so that signal arrives in assigned time slot at hub sector. Intercell uplink cochannel interference is approximated by full-time transmission from worst-case terminal interference from neighboring cell.

at different locations using assigned TS that are usually not fixed assignments but made as the traffic dictates. It is therefore not feasible to design the network, assuming the interference on a given uplink time slot will be originating from a particular terminal.

If the uplink radio channel assignment, but not time slot assignment, is known for terminals in a neighboring cell, the uplink interference on that channel could be calculated by averaging the interference from each terminal on that channel in each neighboring cell. This requires that the path loss from each terminal in each neighboring cell to the victim hub receiver be calculated as

$$I = \frac{1}{N_1}\sum_{i=1}^{N_1} P_{\mathrm{i}} + \frac{1}{N_2}\sum_{i=1}^{N_2} P_{\mathrm{i}} + \cdots + \frac{1}{N_m}\sum_{i=1}^{N_m} P_{\mathrm{i}} \tag{8.9}$$

where $N_{1\cdots m}$ are the number of TDMA terminals in each of the neighboring hubs 1 through m using the uplink radio channel under study. The power of each of the interference terms P_{i} are calculated using (8.6).

The weakness of this approach is that the terminal with the highest interfering power P_{i} conceivably could be actively transmitting for many or all of the channel TS, so worst-case interference essentially exists all the time. The resulting net interference is larger than the average. The conservative approach, therefore, is to use the interference from the worst-case terminal in each neighboring cell and assume that this terminal is actively transmitting during all TS.

$$I = P_{1,\max} + P_{2,\max} + \cdots + P_{m,\max} \tag{8.10}$$

where $P_{1,\max}$ is the uplink interfering power from the worst-case terminal operating in neighboring cell 1.

Interference in TDMA systems can be further mitigated by applying a technique known as *frequency hopping* (FH). FH is one method for creating a spread spectrum communications system that will be discussed in a later section. Its improvement value for TDMA

systems can be appreciated from the foregoing discussion of interference levels on the uplink, and by extension, on the downlink. If the TDMA system uses a set of TS taken from different radio channels, the average interfering power on any given radio channel is reduced. FH with TDMA is an option that is available with GSM cellular systems that can potentially reduce the average interference levels by 2 to 3 dB.

8.4 CODE DIVISION MULTIPLE ACCESS (CDMA)

CDMA uses spread spectrum techniques and their inherent interference immunity to achieve access to the spectrum by multiple users. Spread spectrum systems take an information data stream, including coding and interleaving, and multiply it by a pseudorandom noise (PN) code sequence at a data rate that is much higher than the rate of the information data stream. This process, in effect, spreads the signal over a bandwidth that is much greater than the bandwidth of the information signal.

By coding the signal with a PN sequence and spreading the signal energy over a broad bandwidth, upon despreading and demodulation, both single carrier (SC) interference from outside the system and other spread spectrum signals from within the system can be effectively rejected. The result is a postdetection SINR that is adequate to achieve a low bit error rate (BER) even though the radio carrier SINR at the receiver antenna terminals is usually less than 0 dB. The bandwidth spreading can be achieved in two ways – frequency hopping FH and direct sequence (DS). Both approaches are discussed below.

Spread spectrum systems have recently seen more widespread interest and use for fixed broadband communications, especially for NLOS systems intended to serve residences and for license-exempt bands such as the IEEE 802.11b band at 2.4 GHz where spread spectrum techniques must be used according to FCC Rules. Note that 802.11b specifies direct sequence spread spectrum (DSSS) but this is not a CDMA system since there is no coordination among the codes used by the various network users. The same codes are used by all users and the channels are shared on a time domain basis [see Carrier Sense Multiple Access (CSMA) in Section 8.6].

CDMA is used for simple point-to-point links in both licensed and license-exempt bands. CDMA technology is also potentially well suited for point-to-multipoint NLOS systems that have system design issues that are quite similar to mobile cellular CDMA systems. With a variety of possible applications for CDMA, and with the differences in the CDMA systems that are now being deployed for fixed broadband wireless, the following discussion of CDMA systems is intended to be a generic treatment of the principles involved. However, certain CDMA system types have emerged at this time as leading contenders for use in NLOS networks such as the 3G UMTS W-CDMA standard, which includes both FDD and TDD implementations. The capabilities of W-CDMA as applied to the particular task of fixed broadband system construction will be covered in a following section.

8.4.1 Frequency-hopping spread spectrum (FHSS)

With frequency-hopping spread spectrum (FHSS), a pseudorandom sequence is used to select one of perhaps thousands of frequencies within the available spread spectrum bandwidth. At any given time only a few of the available frequencies are being modulated by

the signal data. Since the spreading bandwidth is much greater than the signal bandwidth, the instantaneous occupied bandwidth of the hopped-to frequency is much less than the total spreading bandwidth. Simple frequency shift keying (FSK) modulation can be used to modulate the carrier. Like Orthogonal Frequency Division Multiplexing (OFDM), because the modulated rate of this carrier is low, it generally is within the coherence bandwidth of the channel and subject to narrowband rather than wideband fading. As such, it provides some immunity to multipath as well as providing frequency diversity. Also like OFDM, it is possible that a given hopped-to frequency will be in a narrowband fade so that the resulting detection of that transmitted symbol (or symbols) results in an error. Therefore, for similar reasons, it is essential to use interleaving and coding to achieve an adequately low error floor.

When the rate of hopping between frequencies occurs at a rate greater than the symbol transmission rate (multiple hops per symbol), the system is referred to as *fast frequency hopping*. If the system hops at a rate that dwells on a frequency for one or more symbol times, the system is referred to as *slow frequency hopping*.

For successful detection, of course, the receiver must have the same random code that determines the hopping sequence so that it can track the hops and correctly detect the signal. Other signals that do not follow this particular pattern are reduced to noise in the detection process. By using multiple random codes, multiple users can share the same spread spectrum channel with each being separately detected among the others by the particular code sequence it uses. In this way, FH-CDMA multiple access to spectrum is accomplished.

Compared to DS systems, FHSSs have not been as widely deployed, although their operation and potential advantage are essentially duals of the advantage available in DS CDMA (DS-CDMA). DS is the approach used in second-generation cellular CDMA systems (IS-95), and in third-generation cellular systems (cdma2000 and UMTS W-CDMA). The wealth of real operational experience and associated large production quantities of the necessary chip sets has made DS-CDMA much more popular than FH-CDMA.

8.4.2 Direct sequence (DS) spread spectrum

In DS-CDMA, the information signal is multiplied by a higher data rate (broader bandwidth) spreading signal so that there are multiple spreading symbols per data symbol. A block diagram of a DS spread spectrum transmitter and receiver is shown in Figure 8.10. The spreading signal is a PN code sequence that is unique to a particular user on a particular hub, thus allowing a separation of the signal at the receiver from other signals in the code domain by once again multiplying by the same unique PN code assigned to that user. The codes can be thought of as a dimension of the spectrum along with frequency and time as shown in Figures 8.1, 8.6, and 8.9. As shown in Figure 8.9, one or more frequency channels may be used for a DS-CDMA system, depending on the data requirements and the available channels. Dividing the data to be transmitted among several CDMA carriers with smaller bandwidth rather than using one CDMA carrier with a wider bandwidth may fit better when a large contiguous block of spectrum is not available.

The transmitted signal is usually binary phase shift keying (BPSK) or QPSK with a symbol rate of R. This signal is multiplied by the binary spreading code with symbol rate

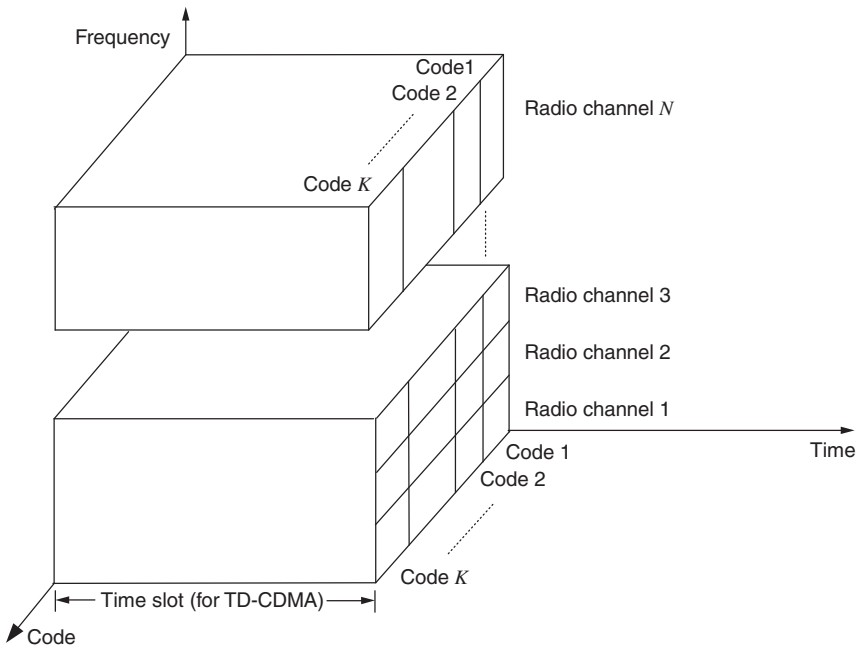

Figure 8.9 Code division multiple access (CDMA) on multiple radio channels.

W, resulting in a transmitted signal symbol rate at the spreading code rate and an occupied bandwidth essentially the same as the spreading code bandwidth (provided $W \gg R$). The transmitted symbols are known as *chips* with the rate of the spreading code called the *chipping rat*e. The ratio of the chipping rate to the data rate, W/R, is called the *spreading factor* (also called *processing gain*). The spreading factor is one fundamental indication of the degree of processing gain in the system, and consequently, the degree of interference rejection (multiuser access) that can be achieved.

As stated in the introduction to this chapter, the core issue with all multiple-access schemes is their ability to operate in the presence of interference. The following sections will present standard interference calculations for CDMA systems. As will be shown, CDMA systems generally rely on careful control of the interference levels from all terminals communicating within a given hub on a radio channel; specifically, best (unbiased) detection of the uplink signals from all terminals connected to a hub is realized when all are received with equal power. This means that CDMA systems must have the ability to control the transmitting power of the remote terminals so that the signals received from terminals close to the hub do not dominate the signals received from more distant terminals. This is the so-called *near–far* problem in CDMA systems. It is dealt with using effective APC that may need a dynamic range of 80 dB or more to equalize the effects of wide path loss differences that can occur with near and far terminals connected to the same hub. The use of advanced techniques such as joint detection (JD) (see Section 8.4.5) can relax the demands on the uplink APC adjustments.

8.4.3 Downlink interference calculations

As with the other multiple-access techniques, the ability of CDMA to support multiple simultaneously communicating terminals depends on the impact of interference from other cells and from the home cell of a given terminal. The composite incoming signal consisting of the desired signal, the interfering signals, and the receiver noise is multiplied by the known spreading signal for the desired signal as shown in Figure 8.10. The signal that was spread by the desired code is recovered and the interfering signals remain as noiselike power across the spreading bandwidth. The SINR can then be calculating by assuming that the interfering signals are additional noise and the desired signal power has been increased by the processing gain or spreading factor W_d/R_d.

The codes used for the downlink and the uplink can be different, so the approach to finding the postdetection SINR is calculated somewhat differently. For the downlink or the forward channel, as it is often referred to in CDMA, the transmissions to all remote units can be synchronized since they are originating at one location and pass over the same channel to each remote unit. Consequently, the spreading codes used are orthogonal codes, sequences of binary pulses that are orthogonal (their inner product is zero). Two types of orthogonal codes are *Walsh functions* used in IS-95 CDMA and orthogonal variable spreading factor (OVSF) codes used in 3G W-CDMA. When the incoming composite signal plus interference is multiplied by the desired signal spreading code, the desired signal is recovered, and ideally, the product of the desired Walsh code with each of the interfering signals is zero. In practice, multipath distortion on the interfering signals to some extent destroys the orthogonality property of the interference, so the interference contribution to the final SINR is not zero.

To establish terminal connections to the network, the hubs typically transmit four kinds of signals – pilot, traffic, paging, and sync. These are referred to as *channels* in the sense of logical rather than physical channels. The pilot signal is detected by the remote terminal to establish a connection to the network, while the traffic channel actually carries the information to be communicated. The ability of the terminal to connect to the network depends on the SINR, or E_c/I_0, of the pilot channel. The quality of the received

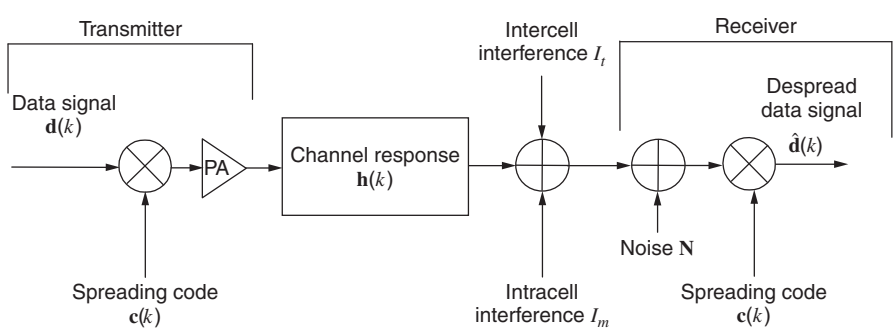

Figure 8.10 Discrete time diagram of a CDMA transmitter and receiver with transmission of block data vector $\mathbf{d}(k)$ for user k.

data depends on the SINR, or more specifically, the E_b/N_0, of the recovered traffic channel data.

8.4.3.1 Downlink pilot channel E_c/I_0

The pilot channel E_c/I_0 as received at a terminal is given by [5]

$$\frac{E_c}{I_0} = \frac{S_R}{I_h + I_o + I_m + I_t + I_n + N_t} \tag{8.11}$$

where

I_h = downlink overhead (sync and paging) power from home hub
I_o = downlink total overhead power (sync and paging) from other hubs
I_m = downlink total traffic channel power from the home hub
I_t = downlink total traffic channel power received from other hubs
I_n = external interference for all other sources such as adjacent channel interference or spurious interference from other transmitters
N_t = receiver thermal noise power
S_R = desired signal pilot channel power received at the remote terminal. This will be some (adjustable) percentage of the total power transmitted by the hub.

The 'home hub' refers to the hub with which the remote terminal is communicating. Since fixed broadband systems do not need to provide for handoff capabilities during transmission, the home hub will normally remain the same hub. The interference quantities listed above are calculated as the sum of the interfering powers as given by (8.6), taking into account the antenna gains and radiated powers in the directions from the hub to the terminal and vice versa. In CDMA systems, the pilot signal is not despread, so E_c/I_0 is usually less than 0 dB (by design, typically -12 to -16 dB for IS-95 CDMA cellular systems).

Like the desired signal S_R, the interference terms in (8.11) are subject to fading due to changing propagation conditions so that E_c/I_0 is also a fading quantity that is properly described in statistical terms. As discussed in preceding chapters, the extent of fading will depend on the propagation environment and the directivity of the antenna used at the remote terminal.

8.4.3.2 Downlink traffic channel E_b/N_0

The downlink traffic channel SINR per information bit, E_b/N_0, can be calculated in a similar fashion to the pilot channel E_c/I_0:

$$\frac{E_b}{N_0} = \frac{S_t}{(1-\alpha)I_h + I_o + (1-\alpha)I_m + I_t + I_n + N_t}\left(\frac{W_d}{R_d}\right) \tag{8.12}$$

where

α = the orthogonality factor (ranging between 0 and 1), which is designed to take into account the lack of perfect orthogonality due to multipath between downlink signals. If perfect orthogonality is preserved, $\alpha = 1$.

W_d = the downlink CDMA chipping rate
R_d = the downlink symbol transmission rate
S_t = desired signal traffic channel power received at the remote terminal for that terminal. The total percentage of hub power dedicated to traffic to all remote terminals is some (adjustable) percentage of the total power transmitted by the hub. The traffic power for a single terminal will normally be the total traffic power divided by the number of terminals being served by the home hub. For CDMA systems with downlink power control, the power allocated to each remote terminal traffic channel will also depend on distance and path loss to that remote terminal.

The definitions of the other interference terms I are the same as given above. The resulting traffic channel E_b/N_0 can be used with the BER equations in Chapter 7 for the data modulation method used to determine the raw error probability. When coding is used, error corrections at the terminal are made to further reduce the net error rate.

8.4.4 Uplink interference calculations

The uplink from the terminal to the hub does not require a pilot channel, although a sync channel may be used if the CDMA system also uses synchronous uplink transmission. For cellular systems, synchronous uplink transmission is problematic owing to the changing mobile locations and the resulting changes in time delays needed to maintain synchronization. The uplink SNR E_b/N_0 is given by

$$\frac{E_b}{N_0} = \frac{S_h}{(1-\beta)(1-\kappa)I'_m + I'_t + I'_n + N_b}\left(\frac{W_u}{R_u}\right)\nu \tag{8.13}$$

where

I'_m = uplink total traffic channel power from the home hub. This is the sum of the power from the other remote terminals connected to this hub and will therefore vary according to the number of these terminals. In general, the APC will adjust the transmitter power for each of these remote terminals so that the received power from each is the same, and is the same as the received power from the subject terminal. This term is also called *multiple-access interference* (MAI).
I'_t = uplink total traffic channel power received from remote terminals served by other hubs. Again, this will vary depending on the number of terminals being served by these other hubs, their locations (path loss to the home hub), and the APC-adjusted transmitter power levels of each of these remote terminals.
I'_n = external interference for all other sources such as adjacent channel interference or spurious interference from other transmitters.
β = the uplink improvement factor (ranging between 0 and 1), which is intended to take into account the reduction in intracell interference (MAI) from the relative degree of preserved orthogonality due to the use of synchronous uplink spreading codes. If perfect orthogonality is preserved, $\beta = 1$. Note that this factor only applies if the CDMA system uses synchronous uplink transmission and orthogonal spreading codes. The IS-95 CDMA system for cellular systems does not provide synchronous uplink transmission, so β is always 0.

κ = the uplink improvement factor (ranging between 0 and 1), which is intended to take into account the reduction in MAI due to the use of JD (see Section 8.4.5). If JD perfectly suppresses MAI, then $\kappa = 1$.
W_u = the uplink CDMA chipping rate
R_u = the uplink symbol transmission rate
S_h = desired signal traffic channel power received at the hub from the remote terminal. This power will vary according to the APC setting for the remote terminal.
ν = a general factor ranging from 0 to 1 that accounts for imperfect power control on the various terminals involved in the E_b/N_0 calculations. For perfect APC action, $\nu = 1$. For imperfect APC, $\nu < 1$ and the resulting E_b/N_0 is degraded.

The value of uplink E_b/N_0 can be used to find the uplink BER using the equations in Chapter 7.

It should be clear from (8.13) that the E_b/N_0 is primarily a function of how many remote terminals are being served by the hub on this frequency (as represented by I'_m) and secondarily by the number of remote terminals connected to neighboring hubs (as represented by I'_t). As the load on the system increases, the power the remote terminal has to transmit to achieve adequate E_b/N_0 increases until the point is reached when the transmit power cannot be increased further. Service to this remote terminal may then be interrupted. This will occur first in the most distant terminals (or those with the highest path loss to the hub), in effect, shrinking the service range of the hub. The dynamically changing coverage area that results from varying cell load is called *cell breathing*.

The increase in interference at the hub is commonly described by the *noise rise* that is essentially the increase in the interference + noise term in the (8.13) from the no-interference case. The noise rise is given by

$$R_N = \frac{(1-\beta)(1-\kappa)I'_m + I'_t + I'_n + N_b}{N_b} \tag{8.14}$$

The hub or cell *load factor* η is a measure of the extent to which the capacity of the cell and hence network is being utilized. The load factor ranging from 0 to 1 is given by

$$\eta = 1 - \left[\frac{N_b}{(1-\beta)(1-\kappa)I'_m + I'_t + I'_n}\right] \tag{8.15}$$

The load factor is normally given as a percentage. For example, given a particular traffic distribution, the system design (hub layout, channel assignments, antenna types, power, etc.) can be configured so that the load factor is 50 to 75%.

The uplink interference level is also a function of the power transmitted by remote terminals in neighboring cells or intercell interference. If the locations of these remote terminals are not known because they are mobile or at fixed but unknown locations, the path loss values cannot be calculated. The interference levels that are needed for summation of I'_t are not explicitly known. In this case the intercell interference is represented by

a *frequency reuse factor* that is usually taken as a fraction of the home cell interference (55%, for example). The frequency reuse factor ξ is defined as

$$\xi = \frac{I_t'}{I_m'} \tag{8.16}$$

Calculating intercell interference using the home cell interference and the reuse factor is a highly approximate approach to finding total interference. However, there are alternatives. For mobile cellular systems in which mobile locations are not known, their location and communication traffic patterns can be simulated through the use of *Monte Carlo* simulation. A Monte Carlo simulation statistically represents the mobiles by guessing at their locations and analyzing the system based on these guesses. The hypothetical locations can be weighted to represent nonuniform traffic distributions throughout the system. This approach will be discussed further in Chapter 11. By using a Monte Carlo simulation for several different random mobile distributions, an overall view of how the system will perform when loaded can be developed.

For fixed broadband systems, the situation is somewhat different in that the location of the remote terminals in the system can be known, at least to the resolution of the address where the system end user has subscribed to the network. Therefore, the opportunity exists to directly calculate the sum of the intercell interference I_t' and therefore achieve a system analysis that is more accurate and that can be dynamically revised as new end users are added.

8.4.4.1 Rake receiver

CDMA terminals use a matched receiver architecture called a *Rake* receiver as illustrated in Figure 8.11. A Rake receiver is essentially a tapped delay line in which the tap delays are equal to one chip symbol time $T_s = 1/W$. As illustrated in Figure 3.6, the received signal includes multipath components to a greater or lesser extent. The objective of a Rake receiver is to isolate multipath energy into resolvable signals spaced at T_s and use those resolved signals as branches in a diversity-combining scheme. The received signal $r(t)$ is sent through a delay line with $L - 1$ delays with delays of T_s. After each delay, the signal at the tap or branch is multiplied by the reference signal $u^*(t)$ (a portion of the desired signal spreading code). This correlation process produces a result that is proportional to the amplitude of the desired signal multipath component. The interfering signals come out of this correlation process as noise. The signal at each Rake tap is then multiplied by a coefficient $C_l(t)$ that depends on the amplitude of the channel energy or SNR in that tap or branch, usually adjusted as a maximal ratio or equal gain combiner. The L branch signals are summed and integrated over the symbol spacing T to yield a decision variable. In this way the Rake receiver collects the multipath energy resulting from the propagation environment and uses it constructively.

Although a Rake receiver is not essential to the operation of CDMA, it does provide useful diversity gain. The amount of diversity gain depends on how multipath-rich the propagation environment is and the chip symbol duration time versus the multipath delay spread. For open environments with no multipath, a Rake receiver provides little diversity improvement. For heavily urban areas, a Rake receiver can provide useful diversity gain.

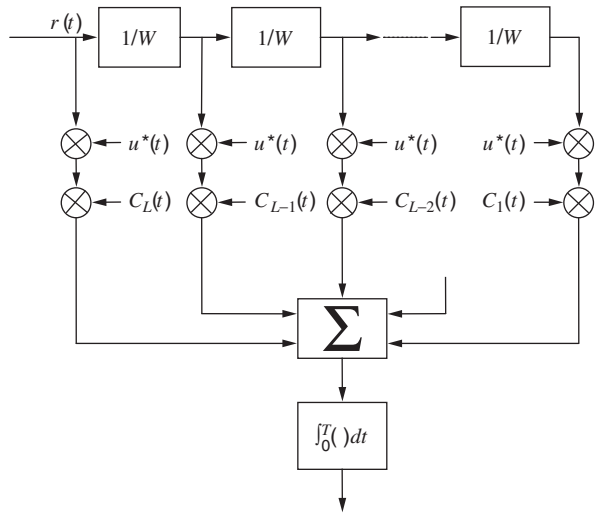

Figure 8.11 Basic rake receiver architecture.

8.4.5 Joint (multiuser) detection

The analysis of the CDMA uplink and downlink E_b/N_0 performance approximated the intracell interference I_m (MAI) and the intercell interference I_t as additional noise in (8.12 and 8.13). For MAI, however, the interference is not random noise but actually consists of signals from other remote terminals served by the home hub. The hub necessarily knows the spreading codes and estimated channel characteristics associated with each of these interfering remote terminals. This information can be exploited to reduce the magnitude of I_m and its impact on E_b/N_0 for the subject user signal. The process of taking advantage of all the uplink signals information arriving at the home hub is called *joint detection* (JD) or *multiuserdetection* (MUD).

The concept of JD as an optimum nonlinear detector for asynchronous CDMA systems was examined in [6] and for synchronous CDMA systems in [7]. These detectors are designed to eliminate MAI but not intersymbol interference (ISI) due to multipath. Since this early work, detectors that suppress both as MAI and ISI have been devised, with the main problem being computational complexity, especially as the number of simultaneous users K increases.

Joint detectors basically function by solving a set of simultaneous linear equations that increase in number as K increases. Joint detectors are therefore most effective (computationally reasonable) when the number of concurrent users is relatively small. It was therefore targeted at TD-CDMA systems such as UMTS terrestrial radio access (UTRA) TDD that are a combination of TDMA and CDMA in which the transmitted data is broken up into time domain frames that are then transmitted on the CDMA spread spectrum channel [8].

Ideally, JD can potentially reduce the term I_m in (8.12 and 8.13) to zero. The capacity improvement (improvement in E_b/N_0) is then given by the reduction in the denominator

of (8.12 or 8.13). If the system is loaded to a reasonable level such that the noise and spurious interference terms, and I_n, are small compared to the system interference, then the upper bound on the capacity improvement ΔC, assuming that JD perfectly suppressed MAI, is given by

$$\Delta C = \frac{I_m + I_t}{I_t} \tag{8.17}$$

As mentioned, for rudimentary CDMA network planning it is often assumed that the intercell interference is some fraction of the intracell interference, for example, $I_t = 0.55 I_m$. The resulting potential improvement factor ΔC from (8.17) is thus 2.8. Unfortunately, the 0.55 factor is just a guess based on simple uniform cell service radii and hexagon cell layout patterns. This value is not supported by detailed simulation of realistic CDMA systems, especially those in highly nonhomogenous propagation environments such as urban areas and microcells where the cell service areas are not contiguous. Consequently, the resulting geographical distributions of interfering remote terminals in neighboring cells are not neatly constrained. The 0.55 factor in reality must be replaced with other values derived from simulations of the actual system. These values could range widely with the resulting improvement factor from (8.17) being equally variable. Whatever capacity improvement does result from JD, this greater capacity can be exploited as support for more users at a given data rate or the same number of uses at higher data rates.

In addition, JD can alleviate the tight accuracy requirements on automatic power control due to the fact that the BER is less dependent on the power settings of the individual remote terminals. The term ν in (8.13) can then be set close to one, also improving the net E_b/N_0.

Although JD is a relatively new technique, it is now seeing use in some commercial deployments. One of the first such deployments using TD-CDMA in the UTRA TDD specification has shown that JD can reduce the intracell interference term I_m to negligible levels [9].

8.4.6 CDMA broadband standards

The modulation and multiple-access schemes for fixed broadband wireless systems discussed here are generally not mandated by regulatory standards bodies leaving the industry free to develop technologies that best meet the perceived needs of the marketplace. One exception to this is third generation (3G) CDMA systems that offer data rates of 2 Mbps (higher in some cases) within the framework of worldwide standards. In particular, the 3G UMTS terrestrial radio access (UTRA) standard provides for FDD and TDD systems with chipping rates of 3.84 Mchips in a 5-MHz bandwidth and data rates up to ~2 Mbps (~4 Mbps for TDD). Basic parameters for these systems are shown in Table 8.2. Both FDD and TDD use a common fixed 15 time slot frame structure with each time slot being 2560 chips long.

For UTRA FDD, clearly a data rate of 2 Mbps cannot be achieved with a 3.84-Mcps chip rate, QPSK, coding overhead, and the lowest spreading factor of 4. A technique known as *multicode* transmission is employed in such cases. With multicode, two or more of the codes available on the hub sector are used to send data to a single remote,

Table 8.2 Basic parameters for 3G CDMA standard systems

	W-CDMA		cdma2000	
Parameter	FDD	TDD	1x EV-DO	3x EV-DO
Chip rate	3.84 Mcps	3.84 Mcps	1.2288 Mcps (downlink) 1.2288 Mcps (uplink)	3 × 1.2288 Mcps (downlink) 1 × 3.6864 Mcps (uplink)
Bandwidth	5 MHz	5 MHz	1.25 MHz	5 MHz
Frame length	10 ms	10 ms	5, 10, and 20 ms	5, 10, and 20 ms
Spreading modulation	QPSK (downlink) QPSK (uplink)	QPSK (downlink) QPSK (uplink)	QPSK (downlink) QPSK (uplink)	QPSK (downlink) QPSK (uplink)
Data modulation	QPSK (downlink) BPSK (uplink)	QPSK (downlink) BPSK (uplink)	QPSK (downlink) BPSK (uplink)	QPSK (downlink) BPSK (uplink)
Spreading factors	4–512 (downlink) 4–256 (uplink)	1, 16 (downlink) 1–16 (uplink)	4–256 (downlink) 4–256 (uplink)	4–256 (downlink) 4–256 (uplink)
Multirate	Variable spreading and multicode	Variable spreading and multicode	Variable spreading and multicode	Variable spreading and multicode
Handoff	Soft handoff	Hard handoff	Soft handoff	Soft handoff
Power control	1.6-kHz update rate	1.6-kHz update rate	0.8-kHz update rate	0.8-kHz update rate
Multiuser detection	No	Yes	No	No

thus multiplying the maximum throughput to a remote terminal by the number of codes used. Of course, if multiple codes are used to communicate with one terminal, fewer codes are available to communicate with other remote terminals at various data rates. Multicodes are also used in UTRA TDD.

For UTRA TDD, the TS are assigned as uplink or downlink slots. The number assigned to a particular direction can be varied in response to the traffic load. One of the 15 TS is used as a beacon channel, a second is used for downlink signaling/control information, and a third is used for uplink signaling/control information. The remaining 12 TS are used for data. At least one slot in each frame must be a downlink or an uplink slot. In order to make the JD interference suppression computationally feasible in UTRA TDD, the maximum number of simultaneous codes that can be transmitted in any time slot is 16. If all codes are devoted to one remote terminal using the multicode technique, the total maximum data rate is (3.84 Mchips/16) $\times$ 16(12/15) = 3.07 Msymbols/s. Since the data modulation is QPSK (2 bits/symbol), the raw maximum bit rate is 6.11 Mbps. The overhead due to coding, midambles data, and so on is about 35%, so the net throughput is about 4 Mbps. The FDD mode allows more simultaneous codes but does not allow for JD for intracell interference suppression. The increase in intracell interference will limit the overall throughput in the hub sector.

The UTRA FDD and TDD standards also anticipate higher chip rates that are multiples of 3.84 Mchips, that is, 7.68 Mchips in 10-MHz bandwidth, 11.52 Mchips in 15-MHz bandwidth, and 15.36 Mchips in 20-MHz bandwidth. Depending on the frequency band and channel allocation constraints where the system is to be deployed, these bandwidths and associated chip and data rates may be viable. The equations in (8.11, 8.12, and 8.13) are generally applicable for system design purposes with arbitrary chip rates, data rates, and spreading factors. Further detailed information on the FDD and TDD modes of W-CDMA systems can be found in published papers [8] and in the voluminous 3GPP specification documents [10].

Table 8.2 also shows the basic parameters for the cdma2000 standards. There are many similarities with W-CDMA but also some distinct differences that make W-CDMA and cdma2000 incompatible. The cdma2000 is more closely related to the 2G- CDMA standard IS-95, which was designed to work in the 1.25-MHz channel bandwidth (one-tenth of the uplink/downlink spectrum width allocated for 870-MHz cellular systems in the United States). The cdma2000 standard has chipping rates and channel bandwidths that continue to fit this channelization scheme. Further details on cdma2000 PHY and MAC layer specifications can be found in the documents of 3GPP2 [11].

All of these CDMA multiple-access technologies have been created to operate in an interference-rich environment. This means that the same frequency channel can be used on adjacent cells and adjacent sectors (a frequency reuse of 1). In comparing the total capacity of various multiple-access technologies and their ability to meet the traffic loads expected in a system service area, the interference immunity properties of a multiple-access technique under *loaded* system conditions must be considered to determine its true usable capacity. This issue is taken up again in Chapters 11 and 12 where specific multipoint systems design and channel assignment methods are discussed.

8.5 SPACE DIVISION MULTIPLE ACCESS (SDMA)

Through the use of high gain fixed directional antennas pointed at the intended service destination, SDMA has been the approach used for many years to share microwave radio channels in the same geographic area. The term SDMA has only recently been applied to make more apparent the parallel purpose of SDMA with TDMA and CDMA in achieving enhanced multiple-access capacity by exploiting all the available dimensions of the radio spectrum resource.

Figure 8.12 shows a sketch of a hub with multiple narrowbeam antennas, each serving a small fraction of the total remote terminals communicating with the hub. Since each antenna is providing separate signals for a set of remotes, it can be regarded as a *hub sector*. With sufficient directionality in the hub sector antennas (see Chapter 6) and in the remote terminal antennas, each hub sector can reuse the same frequencies for most terminals without the need for TDMA or CDMA schemes. The processing gain (relative interference rejection) realized is simply the antenna beamwidth at which sufficient suppression is achieved to prevent interference to a remote terminal on an adjacent hub sector so that its SIR is adequate to provide the service-required BER. For example, if the required SIR is 10 dB and the antenna has a 10 dB beamwidth of Φ, then the number of sectors M that can be situated at the hub is

$$M = \frac{360}{\Phi} \tag{8.18}$$

If the 10-dB beamwidth is 10°, M is 36 or a processing gain of about 15 dB without utilizing additional bandwidth as with CDMA or time-sharing the channel as with TDMA.

As the beamwidth of each sector radiator decreases so that it is only serving a single remote terminal, the processing gain increases still further. The simple LMDS antenna

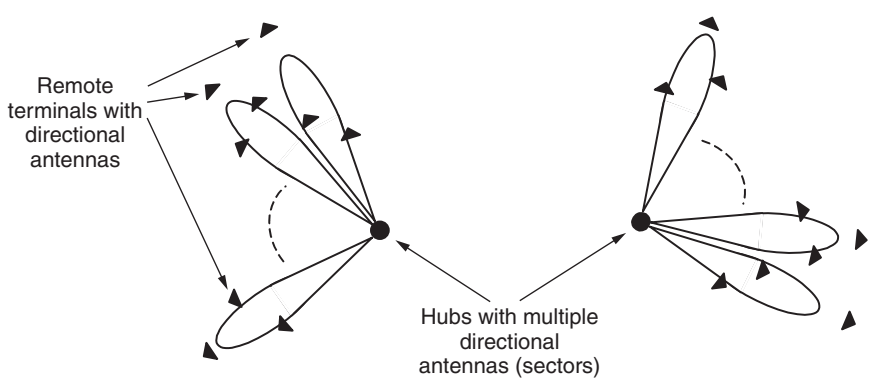

Figure 8.12 Space division multiple access (SDMA) using directional antennas in a point-to-multipoint system to achieve separation of transmissions to/from remote terminals using directional antennas.

whose radiation pattern is shown in Figure 6.12 has a 10-dB beamwidth of about 5°, increasing the processing gain to over 18 dB. The limit of this approach is perhaps represented by free space optics (FSO) in which the laser has a typical beamwidth of about 0.34°, resulting in a processing gain of 30 dB over the omnidirectional case.

These simple calculations are for the 2D horizontal plane (azimuth) case. If the antenna has directionality in the vertical plane as well, and the remote terminals are also distributed vertically as they are in a fixed broadband system in a densely urban area, the vertical plan directionality would make the SDMA processing gain even greater. From these examples, it is obvious that the potential capacity increases available with SDMA are substantial without increasing bandwidth or partitioning transmission time.

Many fixed broadband hubs deployed in the 24-GHz, 28-GHz, or 38-GHz bands are actually collections of point-to-point links using high gain antennas on each end. The 'hub' is simply a building roof or tower top where a portal to an optical fiber or other high-speed backbone link is available, and thus becomes the common terminal for a collection of point-to-point links. Since each antenna is directed to serve a single customer, these antennas are not really hub sector antennas.

The fixed sector antenna approach has been implemented in cellular radio systems as the so-called *switched beam* technology in which the signals from fixed narrow beam antennas are compared to find which has the best signal from the remote terminal (see Section 6.6). For mobile systems, this comparison is periodically done so that the mobile can essentially be tracked as it moves. Suppression of interference from other remote units is basically whatever incidental suppression is provided, given the fixed orientation of the antenna. The fixed beam approach can be replaced with adaptively beam-steered antennas that seek to maximize the signal level from the desired remote terminal. A steered antenna does not explicitly seek to maximize the SIR by suppressing interference as well.

The adaptive or smart antennas described in Section 6.6 use adaptive beam-forming to create a gain maximum toward the desired remote terminal while simultaneously forming pattern gain minima toward cochannel interference sources, thus maximizing SIR. From an achieved SIR and capacity perspective, this action can be modeled as the fixed narrow beam antenna continuously pointed at the other end of the link.

For fixed broadband systems, the problem of adapting the antenna to point at a moving target is eliminated compared to cellular systems. Nonetheless, the propagation environment is constantly changing, even from highly localized effects such as persons moving in a room around the antenna; so to achieve the processing gains possible with SDMA, the antennas must still be capable of adapting to this changing environment.

If the antenna used on the remote terminal end of each link is a low gain antenna that illuminates a wide extent of the propagation environment, the angular spread of the signal arriving at the hub may not be a single direction but span a range of directions due to scattering. The interference source can be scattered in a similar way. The angular spread of the desired and the interfering signals inhibits the ability of the adaptive antenna to maximize the SIR. However, when compared to a mobile system, it is possible to exploit the relatively lower variability of the angular spread in a fixed broadband system with longer sampling periods in the adaptive antennas to arrive at the best SIR results.

The number of discrete interferers that can be suppressed using an adaptive antenna is equal to $N - 1$, where N is the number of antenna elements. For the simple adaptive antennas being designed for use in remote terminals on NLOS system, only 2 or 3 elements are used. If 2 or more strong interferers exist, such simple adaptive antennas may not be able to suppress all the interference and the SIR will not be adequate. This scenario can only be investigated in real terms when a system is fully deployed and fully loaded with remote terminals. It has been the experience with cellular systems that they are designed with an over-abundance of downlink signals in certain service areas (called *pilot pollution* in CDMA systems). Similar design errors in a SDMA system, in which adaptive antenna directionality was relied upon to achieved adequate SIR, could lead to more strong downlink signals than the remote terminal can successfully suppress. Avoiding this situation is an important aspect of designing a fixed broadband system that relies on SDMA and adaptive antennas to achieve multiple access.

8.6 CARRIER SENSE MULTIPLE ACCESS (CSMA)

A straightforward means of achieving multiple access on a radio channel is to establish a protocol in which each remote terminal monitors the radio channel and does not attempt to transmit any data until it detects that the channel is currently idle (no other terminal is attempting to transmit). This approach is quite general and can be applied to form a wide variety of 'ad hoc' wireless networks consisting of user terminals and network access points. This technique is usually called carrier sense multiple access (CSMA) because the network nodes 'sense' the presence of other radio carriers before transmitting.

Of course there are limitations to this approach. The primary problem is that two or more terminals may attempt to transmit at the same time because neither sensed the other first, such that some part of their transmissions overlap. The resulting interference will cause the transmission to be received with errors. When such collisions are possible, the network can remedy the situation in several ways:

- *CSMA/CD*: The 'CD' appended to CSMA stands for *collision detection*. When a terminal detects a transmission collision, it terminates the transmission and waits for an open time to begin transmitting. This approach is used with wired Ethernet networks and is viable on other networks when each terminal can detect the transmissions of every other terminal.
- *CSMA/CA*: The 'CA' stands for *collision avoidance*. This approach is appropriate for wireless networks when all the terminals cannot necessarily detect the transmissions of other terminals. Collisions are avoided by a terminal first inquiring of the receiving terminal if it is acceptable to transmit using a short request to send (RTS) message. If the receive terminal sees a clear channel, it responds with a Clear to Send (CTS) message that also tells others terminals (within detection range) that the channel is busy. When a transmission is received without error, the receiving terminal can send an acknowledgement (ACK) of correct reception or, conversely, a NACK message (negative acknowledgement). This is the basic (though not the only)

approach for CSMA in the IEEE 802.11a and 802.11b wireless network standards. The timing and time-out restrictions in these standards also impose fundamental constraints on their usable range.

Unlike the other multiple-access techniques discussed above, CSMA does not provide for simultaneous use of the radio channels, although it is somewhat like TDMA without reserved time slots and thus, without reserved spectrum capacity for active users.

The drawback of CSMA is that terminals must wait for an open window in which to transmit. That wait gets longer as the number of connected terminals increases. The probability of collisions is directly related to the propagation delay between terminals and the detection delay at a given terminal, and, of course, the number of terminals. As the probability of collision increases, and retransmission (sometimes multiple retransmissions) is needed to ensure correct data receipt, the net data throughput of the network decreases dramatically and the turnaround delay experienced by the terminal user increases.

8.7 MULTIPLE ACCESS WITH OFDM

OFDM (or COFDM in normal use) as described in Chapter 7 by itself is not a multiple-access method. It is subject to cochannel interference from other users in the network that will result in detection errors depending on the type of modulation used for each OFDM subcarrier.

OFDM is used in *single frequency networks* (SFN) in which cochannel interference is anticipated and planned. Correct operation of an OFDM SFN requires exact synchronization of all transmitters in the network as well as appropriate selection of the guard interval so that echo signals (interference) from other transmitters occur within the guard interval and are thus ignored, or occur within the same transmitted symbol time so they can be used constructively in the receiver to improve system performance. However, an excessive guard interval to successfully achieve SFN operation represents idle transmission time that directly subtracts from the overall capacity of the channel. Also, such systems are really not multiple-access systems because the transmitted information from each station in the network is the same.

Since OFDM is only a modulation type, some other process must be included for it to be used with efficient and competitive capacity in a multiple-access network. Five main approaches have emerged to accomplish multiple access using OFDM as the modulation type:

- multicarrier CDMA
- orthogonal frequency division multiple access (OFDMA)
- OFDM with TDMA
- OFDM with CSMA/CA
- OFDM with SDMA.

Each of these will be discussed below.

8.7.1 Multicarrier CDMA (MC-CDMA)

Despite its name, multicarrier CDMA is actually a method of employing OFDM to multiple-access scenarios. With MC-CDMA, the spreading code is used to spread the data symbols across multiple OFDM carriers. The fraction of the data symbol corresponding to a chip in CDMA is transmitted via a single subcarrier. The entire data symbol is therefore transmitted at one time on the entire ensemble of OFDM subcarriers. This necessarily means that the symbol rate on any individual subcarrier is the same as the data rate. Successive data symbols are transmitted in the same way.

An important aspect of OFDM is that the symbol rate on the individual subcarriers must be slow enough, and hence their occupied spectrum must be narrow enough, such that frequency-selective fading does not occur across the spectrum of any subcarrier. If this is not the case, with MC-CDMA the data stream must be further subdivided using a serial-to-parallel converter, with each data symbol from the output of the converter further subdivided to modulate a set of OFDM subcarriers in the usual way as described in Section 7.3.3. A block diagram of this process is shown in Figure 8.13.

The number of subcarriers needed for MC-CDMA depends on the extent to which the serial-to-parallel converter must group data blocks for transmission. If the converter sets up P output bits for transmission, the number of subcarriers required is

$$N = PK_{MC} \tag{8.19}$$

where K_{MC} is the spreading code length. The total occupied bandwidth BW_{MC} is given by

$$BW_{MC} = \frac{(PK_{MC} - 1)}{(T_S - T_G)} + \frac{2}{T_S} \tag{8.20}$$

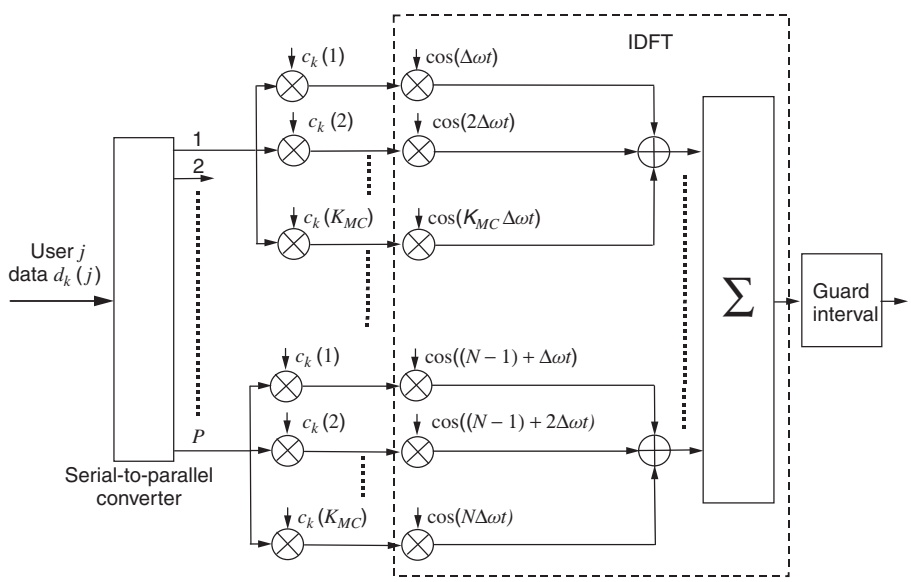

Figure 8.13 Basic block diagram of an MC-CDMA transmitter for user j.

where T_S is the subcarrier symbol duration and T_G is the guard interval. The subcarrier spacing is then $\Delta\omega = 2\pi\,\Delta f = 2\pi/(T_S + T_G)$.

An analysis of MC-CDMA as compared to DS-CDMA is presented in [12] for several different system designs and multipath scenarios. This comparison arrives at the conclusion that MC-CDMA has no major advantage over DS-CDMA in terms of required bandwidth to achieve the same BER performance. Its ability to reject interference and therefore provide capacity in a multiple-access system design is also the same as DS-CDMA. Like other OFDM systems, however, MC-CDMA does place greater linearity requirements on the power amplifier compared to DS-CDMA, as well as subcarrier stability constraints to avoid intercarrier interference (ICI). Using SC modulation, there obviously are no ICI problems with DS-CDMA.

While MC-CDMA has undergone significant analysis by a number of researchers (see the reference list in [12]), at present it is not a technology that is being pursued for use in fixed broadband wireless system deployment.

8.7.2 Orthogonal frequency division multiple access (OFDMA)

OFDMA is a logical extension of OFDM in which a subset of the available subcarriers is used to carry the data for a given remote terminal. The subcarrier universe is that established via OFDM, but the subset of subcarriers assigned to each user can be (randomly) distributed among the universe of subcarriers. This makes it somewhat similar to FDMA if a user's data were subdivided among multiple subcarriers; however, because of the orthogonal relationship of the carriers in OFDM, it is easier to filter one subcarrier from another, so guard bands and tight filtering as with FDMA are not needed.

From the use of multiple subcarriers, it is an obvious step to further randomize the channel usage and average the interference potential by employing FH among the OFDM subcarriers. In FH-CDMA, the frequencies used in the hopping sequence can be arbitrary as long as the receiver knows what the frequencies are and the order in which they are accessed. With OFDMA, the frequencies are established by the OFDM modulation with the FH sequence used to distinguish the signal of one terminal from another. The multiple-access concept can be further extended to subdivide the spectrum in time. Individual remote terminals are assigned TS as in TDMA during which they are permitted to access the subcarriers to transmit data via their uniquely assigned coded hopping sequence.

As with any FH system, the hopping patterns can themselves be orthogonal; that is, for N frequencies in the systems, N orthogonal hopping patterns can be constructed. OFDMA is one of the modulation/multiple access techniques specified in the draft of IEEE 802.16a (see Section 8.7.6). Because all the subcarriers used on a hub for communication with various remotes are orthogonal, intracell interference is eliminated, subject to phase/frequency distortions from the multipath. The only interference source will be from remotes and hub sector transmitters in other cells. This interference will experience spread spectrum interference averaging over the entire channel bandwidth as normally seen with spread spectrum (SS) systems in general and CDMA in particular.

8.7.3 OFDM with TDMA

Any of the modulation types described in Chapter 7 can be used with TDMA by simply defining a TDMA frame structure with TS and using the selected modulation during each

time slot to communicate with separate remote terminals. Using TDMA with OFDM is one of the multiple access/modulation methods included in the draft of the IEEE 802.16a standard (see Section 8.7.6).

8.7.4 OFDM with CSMA/CA (IEEE 802.11a)

This combination of technologies is included here because it is the combination selected for IEEE 802.11a systems operating in the 5-GHz license-exempt band in the United States and Europe. Both OFDM and CDMA/CA were discussed in earlier sections.

The 802.11a standard currently provides for up to 12 nonoverlapping 20-MHz-wide channels, each with a raw data rate of up to 54 Mbps at short ranges. The allowed transmitter power output ranges from 50 to 250 mW for the 8 low band channels intended for indoor communications, and up to 1000 mW on the upper 4 channels intended for outdoor building-to-building communications. Each 20-MHz channel is occupied by 52 OFDM subcarriers resulting in a carrier spacing of about 300 kHz. To achieve the maximum transmission a rate of 54 Mbps, each of the OFDM subcarriers is modulated with a 64QAM signal as described in Chapter 7.

As with 2.4-GHz IEEE 802.11b that uses a DS spread spectrum air interface (not CDMA), the multiple-access scheme in IEEE 802.11a still relies on the relatively inefficient CDMA/CA approach. For a large number of remote terminals attempting to use the network, the net data throughput will decrease and the network response delay will increase.

8.7.5 OFDM with SDMA

Some systems currently being planned for NLOS operation in the licensed bands at 2.5 to 2.7 GHz and 3.5 GHz use a combination of OFDM and adaptive antennas to accomplish multiple access. This combination provides good multipath immunity and interference rejection properties that will potentially allow multiple users to operate on the same frequency on the same hub sector. The relative ability of OFDM to handle cochannel interference as already discussed is modest, so the multiple-access strengths of the scheme rely solely on the ability of the adaptive antennas to enhance the desired signal and suppress interference. As mentioned earlier, several strong interferers will present the greatest challenge to OFDM and SDMA systems. In such cases, with multiple strong interfering signals close by, intelligent assignment of different frequency channels is an appropriate solution. With an NLOS system and remote terminals intended to be self-installed by the end user, the frequency selection process must also be automatic, even with the potential close-by interferers not yet in operation. Frequency assignment strategies of this type are discussed in Chapter 12.

8.7.6 OFDM multiple-access standards

There are at least two standards in which OFDM is currently specified as a modulation type with associated multiple-access methods. The two most relevant to current fixed broadband wireless system design are the IEEE 802.11a standard for the 5- to 6-GHz band

Table 8.3 Draft 802.16a OFDMA parameters

Parameter	Value
FFT size, N_{FFT}	2048, 4096
Bandwidth, BW	Depends on band and country
Sampling frequency, F_S	$BW \cdot (8/7)$
Subcarrier spacing, Δf	F_S/N_{FFT}
Useful time, T_b	$1/\Delta f$
Cyclic prefix (CP) time, T_g	Based on T_b and allowed ratios
OFDM symbol time, T_S	$T_b + T_g$
Ratio, T_g/T_b	1/32, 1/16, 1/8, and 1/4
Sample time	$1/F_S$
Number of downlink/uplink channels	32
Number of data subcarriers per channel	48

and IEEE 802.16a standard for the 2- to 11-GHz band. The IEEE 802.11a is described in Section 8.7.4.

The draft IEEE 802.16a standard [13] for wireless access systems in the 2- to 11-GHz range is intended for much longer range metropolitan area networks (MAN) rather than local area networks (LAN), which is the target of IEEE 802.11a. The IEEE 802.16a standard has specifications for both licensed and license-exempt bands. IEEE 802.16a specifies the use of three different modulation/multiple access techniques:

- single carrier (SC) with TDMA;
- OFDM with TDMA: The fast Fourier transform (FFT) size can be 256 or 512, resulting in 212 or 424 active subcarriers, respectively;
- OFDMA with FFT sizes of 2048 or 4096, with 1696 or 3392 active subcarriers, respectively. For the 2048 case, the subcarriers are divided into 32 channels of 48 subcarriers each, with the remaining subcarriers used as pilots and other overhead functions.

Given the large number of bands in which the 802.16a standard can be applied worldwide, there is no fixed channel bandwidth. Instead, the OFDM subcarrier frequency spacings are scaled for the available channel bandwidth. Some of the basic parameters are shown in Table 8.3. Both FDD and TDD duplex methods are permitted by IEEE 802.16a. Given the various options available in this standard, transmissions contain preamble information that instructs the remote units as to the technology being used.

8.8 DUPLEXING METHODS

As discussed in the introduction to this chapter, duplexing is the mechanism by which two-way communication is accomplished between two points in a single link system or between a remote terminal and the network hub or backbone in point-to-multipoint

networks. The duplexing method chosen for a system is essentially independent of the modulation type and the multiple-access-technology. The main issues that influence the duplexing method selection are the nature of the available spectrum allocations and the nature of the traffic that is to be communicated.

The two duplexing methods primarily used in fixed broadband wireless are FDD and TDD. Both are discussed in the following sections along with the lesser-used methods of polarization duplexing and space duplexing.

8.8.1 Frequency division duplexing (FDD)

Frequency division duplexing is the method most widely used in traditional fixed broadband point-to-point and point–to-multipoint systems. A drawing illustrating FDD is shown in Figure 8.14. The primary parameters with FDD are the channel bandwidths themselves for downlink and uplink and the duplex frequency separation between pairs of downlink and uplink frequencies. The downlink and uplink bandwidths are chosen to accommodate the required system data rate and modulation scheme, or they may also be mandated by administrative rules as in the case of first- and second-generation cellular systems. The total duplex bandwidth available to a user is $D_n + U_n$. If asymmetrical downlink and uplink traffic volume is anticipated, D_n does not necessarily have to be the same as U_n, although in conventional systems it is typical for the downlink and the uplink bandwidths to be equal.

The duplex frequency separation is chosen largely to make the design of duplexing waveguides or circuitry at a terminal practical. For a full duplex system, a terminal will be sending and receiving simultaneously so that both the downlink and the uplink channels are active. If a single antenna is used, which is the economical approach, the device known as a *duplexer* at the remote terminal must separate the signals being received on D_n from those being transmitted on U_n. The design of duplex filters to accomplish this task becomes more difficult and problematic as Δf decreases. Since the uplink and the downlink signals are different, signal leakage of one into the other represents interference and

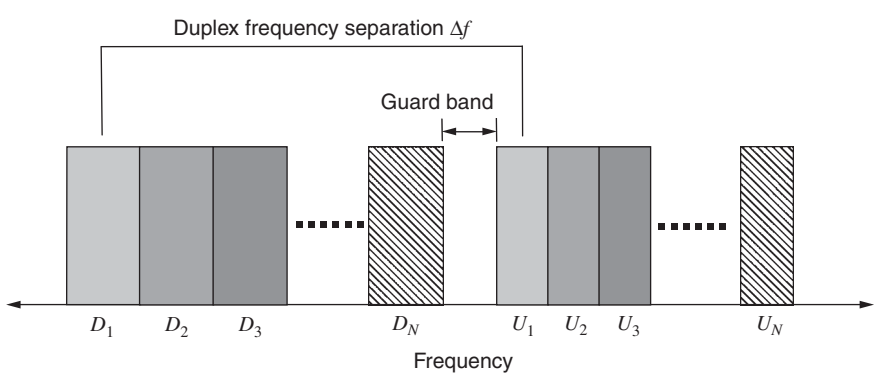

Figure 8.14 Frequency division duplexing (FDD) with downlink channels D_n and uplink channels U_n with duplex frequency separation Δf.

correspondingly lowers the fade margin. A typical minimum duplex frequency separation is 5% of the carrier frequency. For the 28-GHz LMDS band, this is about 1.4 GHz.

Anything that can increase the isolation between D_n and U_n will relax the filtering requirements of the duplexer. For microwave systems with fixed high gain antennas, a common approach for improving downlink–uplink isolation is to use one polarization for the downlink and the orthogonal polarization for the uplink. As illustrated in Figure 6.12, in the main lobe of antennas typically used in such systems, the order of 30 dB of isolation can be achieved between orthogonal polarizations. In effect, this is *polarization duplexing*. The required duplex frequency separation Δf can be significantly reduced with effective polarization duplexing.

From a spectrum efficiency standpoint, the main drawback to FDD is that it reserves spectrum in D_n or U_n that may be underutilized much of the time. This is especially true when the traffic is highly asymmetrical. The usual situation considered is the 'web-browsing' application in which the downlink transmission is a relatively high volume of data traffic including web pages, images, or video or audio files, while the uplink transmission consists simply of a series of mouse clicks or key strokes – relatively low traffic volume. If a full uplink channel is reserved for this terminal, it could well be idle most of the time, which is an inefficient use of the spectrum. This is the argument most often put forward in favor of TDD discussed in the next section.

Other data applications such as e-mail represent much more symmetrical downlink–uplink traffic volume distribution, and the potential benefits of TDD are not quite so apparent. Also, for traditional two-way point-to-point microwave links that carry high volume data traffic in both directions, or multiplexed voice traffic with a fixed number of channels in each direction (even if idle), TDD can have drawbacks. Such point-to-point links are commonly used in cellular backhaul (network interconnect) applications or corporate interbuilding or intercomplex connections. For such situations, FDD is the better approach for duplexing.

8.8.2 Time division duplexing (TDD)

Time division duplexing (TDD) uses the same frequency channel for both downlink and uplink transmission. For conventional analog wireless systems, the TDD operation is called *simplex*. In a typical simplex analog wireless system, the network user listens for an idle channel before transmitting a message. The user then monitors this same frequency for a response to the transmitted message. What distinguishes TDD from simplex systems is that modern digital hardware makes it possible to rapidly switch the channel from transmitting to receiving information, and interleave (in time) packets of transmitted and received data such that from the end user's viewpoint, the channel is perceived to be a full duplex channel that can transmit and receive information 'simultaneously'. A drawing illustrating TDD is shown in Figure 8.15. A guard time is inserted for transitions from transmit to receive mode, and vice versa, between TS.

The main advantage of TDD is that the time occupancy of the channel can be adjusted to accommodate the asymmetrical traffic volume on either the uplink or the downlink. The reduction in channel idle time leads to greater efficiency than FDD in which some channel idle time would occur with asymmetrical traffic volume. For an asymmetry factor of 2, the

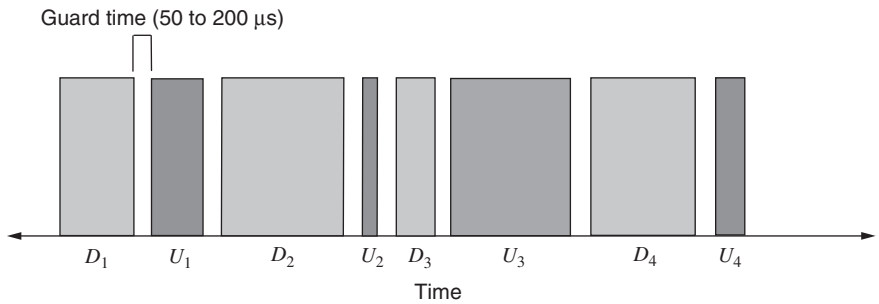

Figure 8.15 Time division duplexing (TDD) with downlink transmission intervals D_n and uplink transmission intervals U_n with guard time. Transmission interval length can vary with adaptive TDD.

increased efficiency is given at about 16% in a recent report. Because of basic overhead information, all TDD systems (like UTRA TDD, for example) use a fixed frame length and TS or 'burst' structure of some sort that becomes the quantum unit for uplink/downlink transmission. The flexibility to adjust the symmetry of the uplink/downlink flow is therefore not unlimited but restricted to possible ratios as discrete values. The limited number of discrete values still provide for considerable range in adjusting the uplink/downlink time slot allocations.

For multiple access or adaptive antenna techniques that use channel estimation, the fact that the uplink and the downlink channels are the same can be used to some advantage. TDD systems are also more flexible in administrative frequency allocations in which only a single rather than a pair of frequency channels are available. Of course, for equivalent full rate downlink and uplink transmission, the bandwidth of the TDD channel must be equal to the bandwidth of the sum of the downlink and uplink channel bandwidths used for FDD.

8.8.2.1 TDD interference calculations

Although TDD has many benefits, one drawback of a TDD system that is rarely pointed out is the doubling of the number of interference sources that must be dealt with as compared to an FDD system. This situation is illustrated in Figure 8.16(a) for FDD with a simple two-hub system with one remote terminal connected to each hub using the same downlink and uplink frequencies f_d and f_u in each case. The interference paths for the FDD interference analysis are shown by dotted lines for the uplink and dashed lines for the downlink. The interference analysis for the downlink (interference to remote terminals) is confined to interference from other system hubs only. Similarly, uplink interference calculations (interference to hubs) are confined to interference from other remote terminals only.

The TDD interference situation is shown in Figure 8.16(b). Because both the hub and remote terminals are transmitting on the channel f_d, both must be considered in the interference calculations at the hub and remote terminal (uplink and downlink) on

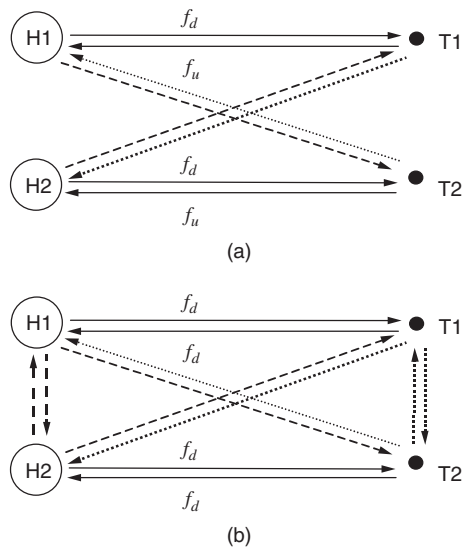

Figure 8.16 (a) FDD interference calculations paths and (b) TDD interference calculation paths.

the neighboring H2-T2 link. While the uplink/downlink time division allocation of the channel is synchronized so that the transmissions do not conflict between a single hub and its remote terminals, this uplink/downlink time division allocation does not generally extend to other hubs and communication with their respective terminals. Synchronizing the two-way uplink/downlink allocation is difficult because the allocation between a hub and remote terminal is traffic-dependent. Synchronizing them all with other links would impose severe uplink–downlink allocation restrictions on those links. Because the uplink–downlink transmission allocations are asynchronous, both the hub and the remote terminal must be considered as interference sources with some reduced average duty cycle. The interference paths for the equivalent two-hub TDD system are shown as dotted and dashed lines in Figure 8.16(b). For UTRA TDD, frame synchronization can assist in mitigating this interference. Some simulations of this problem for UTRA TDD are found in [11].

For systems using SDMA that rely on adaptive antenna technology for interference suppression, the increase in the number of potential interference source locations means these antennas have more interference sources they must suppress. For remote terminals with modest adaptive antennas made up of only a few elements, the increase in interference source locations (and corresponding different arrival angles) will increase the difficulty of adjusting the antenna to suppress interference and achieve the desired SIR.

8.9 CAPACITY

A measure of the success of any multiple-access technique is the overall capacity it can achieve in terms of bps/Hz/km^2 throughput in a two-dimensional service area. Fixed

broadband systems in densely urban areas actually operate in a three-dimensional propagation environment, which can be exploited using SDMA or fixed broadband microwave systems using antennas with high gains in both the azimuth and elevation planes. Very narrow beam FSO technology can best exploit the 3D spectrum space. Capacity for such systems should really be recognized in three dimensions, or bps/Hz/km^3. While this wireless 'box of beams' is an accurate generalized view of a fixed broadband network deployment, for simplicity, capacity will be considered here in the more globally applicable two-dimensional framework.

Unfortunately, the raw capacity metric bps/Hz/km^2 is not very useful for assessing the quality of the communications process under a variety of single-user, multiuser, noise-limited, and interference-limited circumstances. Capacity can be assessed in a variety of ways depending on how that assessment will be used. A number of capacity calculation approaches are discussed below including

- Shannon theoretical channel capacity limit
- capacity in interference-limited, multiuser systems
- user capacity
- commercial capacity.

Capacity figures are often used to make relative comparisons of the success or strength of one technology versus another, but sometimes these comparisons are influenced by other nonengineering issues that are not relevant to a system operator trying to select a technology to meet service objectives and business plan milestones. With that in mind, the following discussion will describe capacity in terms of what a communication system operator can offer customers by way of data throughput, reliability, and service quality.

8.9.1 Shannon theoretical channel capacity

A presentation of Shannon's theory for the upper bound of channel capacity in bits per second for a channel corrupted by additive white Gaussian noise (AWGN) can be found in any modern communications engineering textbook, for example, [14,15]. The Shannon capacity bound is a rate at which information can be transmitted with arbitrarily small error rate (including zero) over an AWGN channel with the use of appropriate coding (which Shannon did not identify). That Shannon information capacity bound is given by

$$C = B_W \log_2 \left(1 + \frac{S}{N}\right) \text{bps} \tag{8.21}$$

where S is the signal power, N is the noise power in the channel, and B_W is the channel bandwidth in hertz. The normalized channel capacity bps/Hz is then

$$\frac{C}{B_W} = \log_2 \left[1 + \frac{E_b}{N_0}\left(\frac{R}{B_W}\right)\right] \text{bps/Hz} \tag{8.22}$$

where N_0 is the noise power density in hertz, E_b is the signal energy per bit, and R is the information rate in bps.

The Shannon capacity limit is not realizable with currently known coding schemes. More importantly, coding schemes that approach the limit require significant lengths and complexity; the resulting long delays tend to infinity as the Shannon limit is approached. While achieving the maximum information capacity from a channel may be desirable, from a user's perspective, infinite delays and the cost associated with high complexity are not. Short of this upper limit, it is possible to identify a transmission rate and particular coding scheme rate or set of rates, that achieve some nonzero (but acceptable) error level.

For most wireless channels, the SNR in (8.21 or 8.22) is a time-varying quantity with a statistical description. The capacity is also similarly varying with time and statistically described. Even though the Shannon capacity is not a very practical measure for wireless system design and planning, it can be a useful benchmark for those engaged in designing coding and multiple adaptive antennas technology for wireless systems.

The Shannon capacity is a channel capacity rather than a system capacity. It relies on some notion of a defined 'channel', which for Shannon in 1947 was better represented by a telephone wire than a complex wireless communications channel. The increased capacity results with multiple input, multiple output (MIMO) systems described in Chapter 6 leads to the view of a 'channel' that is not even a distinct entity but rather a continuum of decoupled transport mechanisms that are enabled in partial, time-varying ways by the complexity and fractured nature of the propagation environment.

8.9.2 Capacity in interference-limited, multiuser systems

Beyond consideration of a single channel, the more generalized capacity of interest is the overall capacity of the network when there are multiple channels and multiple users. More directly, the network capacity is the number of users whose communication objectives can be 'simultaneously' met, with the caveats on 'simultaneously' previously discussed in the introduction to this chapter. This approach to calculating capacity has been used extensively for cellular system technology comparisons. The basic weakness is that analytical studies require several assumptions, particularly, interference levels from other parts of the system and the channel response characteristics. The traditional hexagon cell layout with assumed cell service radii with associated intercell interference and reuse distances is only theoretical – no real cellular system has this topology. As the propagation environment gets less homogeneous, and directional antennas are used, the interference levels leading to BER calculations are not uniform. The system capacity thus becomes less dependent on the technology and more dependent on the environment where the technology is deployed. Environment-specific simulations with full-featured propagation, channel, and hardware models are the appropriate approach to assessing the potential capacity of competing technologies.

The factors in a multiuser system that affect the system capacity are as follows:

- Required net error rate for a given application after decoding and forward error correction (FEC)
- Modulation efficiency (bps/Hz)
- Coding overhead

- Multiple-access technique – FDMA, TDMA, CDMA, SDMA, and OFDMA
- Diversity (time, space, frequency, code, etc.)
- Overhead for guard bands, guard times, and so on
- Hard or soft handoff network overhead requirements
- Directional antennas
- Automatic power control (APC) on both uplink and downlink
- User service duty cycle or packet traffic statistics (voice activity factor in voice systems)
- System intelligence in terms of dynamic resource assignments to respond to changing service demand. Dynamic channel assignment (DCA) and dynamic packet assignment (DPA) are two examples of system intelligence that affects capacity.

As an example, consider a 1-MHz channel using 16QAM modulation. Theoretically 16QAM has an efficiency of 4 bps/Hz, but with coding overhead and imperfect filtering, the net actual efficiency is typically around 3.0 bps/Hz. In a channel with a 1-MHz bandwidth, the data rate furnished to the user will be 3.0 Mbps. If the channel is time-shared using TDMA among 10 users, and TDMA overhead is ignored, the data rate to each user is 300 kbps. If the system has 10 such channels using a single hub, the total hub capacity is 30 Mbps.

The above calculation assumes that the SNR at the user terminal is adequate to achieve an acceptable error rate after decoding and error correction. For a given transmitter power, receiver noise figure, and antenna gains, this implies there is a limit to the range or link pathloss at which service can be furnished. Choosing a more efficient modulation scheme such as 64QAM with 6 bps/Hz will increase the total data rate available in a channel but will reduce the transmission range over which this data rate is available because a higher SNR is required to achieve the required error rate. The more efficient modulation is also more susceptible to errors from interference as shown in Chapter 7, so the ability to reuse frequencies on nearby hubs is reduced, thus reducing the overall *network capacity*.

A multihub hypothetical hexagon system layout is shown in Figure 8.17. For a multihub system, each with multiple users, the performance of each link on each of the 10 channels is now degraded by intercell interference. The extent of interference will depend on the distance and resulting pathloss between the receivers (both hub and remote terminal) in a victim cell and the transmitters (both hub and remote terminal) in neighboring cells. The interference level will also depend on the directional antenna properties of all transmitters and receivers, and whether APC is used on transmitters. If the locations of all transmitters and receivers are explicitly known along with the characteristics of the antennas and RF equipment, a pathloss model can be used to find the interference level from every transmitter to every receiver. The capacity of the 10 channels used by the receivers can then be reassessed in light of the interference + noise level. A similar procedure can be used for each cell and the total capacity of the network explicitly calculated. Of course, when anything changes (more terminals, different modulations, etc.) the system capacity must be recalculated. Chapter 12 includes the mathematical details of the capacity calculation.

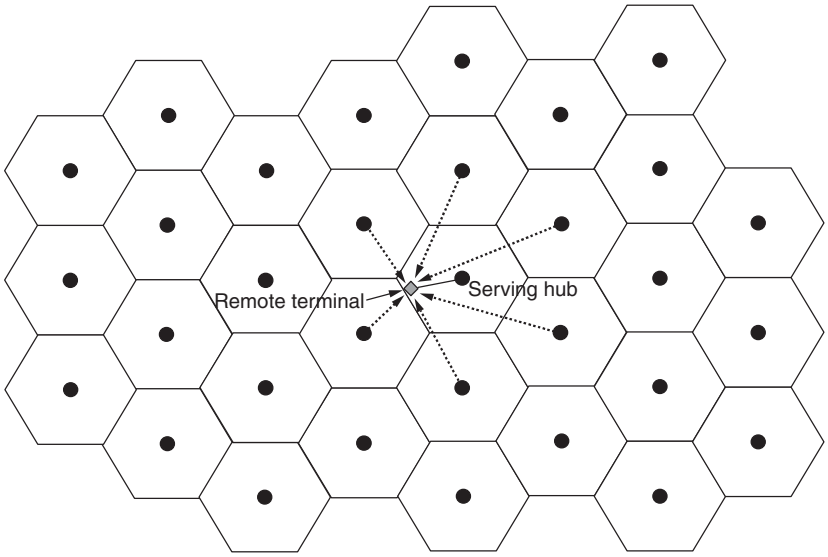

Figure 8.17 Hypothetical hexagon grid cell layout with interference paths to remote terminal from closest (first 'ring') interfering hubs shown as dotted lines.

The point of describing this example system is to demonstrate how dependent the system capacity is on all the bullet points listed above. In simple analytical analyses of networks, many of the direct calculations described above are replaced by assumptions. For example, the service range for a given modulation type is usually assumed to be a fixed distance from the hub, and the interference from other cells is summarized by a single number that is proportional to its distance from the victim cell. Even though these assumptions cloak the details and complexity of the problem and thus lowers the quality of the capacity estimate, this approach at least provides a simple, easily understood 'ballpark' capacity number.

In Chapter 11, specific link budget calculations will be presented that quantitatively take into account the factors cited above. Comprehensive, iterative use of such calculations, used in the context of wireless system planning software, can provide much more realistic estimates of the capacity of particular system design by taking into account all the particular parameters of that design.

As a general principle, network capacity will increase as the cell service area size decreases. This is simply a result of shifting the capacity from the wireless radio resource to a larger number of network access point resources. However, the infrastructure cost associated with this shift is significant, and includes the infrastructure access points themselves (including towers, etc.) and the cost of connecting these access points to the optical fiber backbone (backhaul). For mobile systems, there is also the issue of handing off service from one access point or hub (cell site sector) to another. Depending on the handoff scheme, some capacity may be sacrificed in maintaining multiple remote-hub connections. This cost trade-off is another aspect of assessing system capacity.

8.9.3 User capacity

The capacity calculations for the system in the preceding section considered service to be delivering a given data rate at a given BER. The concept of capacity can be extended to calculations of how many users are receiving 'acceptable' service that include several other QoS parameters. Such additional parameters include

- connection time overhead (how fast does it connect)
- turnaround delay (latency)
- jitter (the variation in latency)
- total uplink and downlink throughput
- throughput variation (service consistency).

A user's perception of the network is based on a qualitative assessment of these characteristics. Bridging the gap from the qualitative user perception to specific engineering design objectives is one of the challenges in modeling network traffic requirements. These aspects of service quality will be discussed in more detail in Chapter 9 on traffic and application modeling.

8.9.4 Commercial capacity

In most cases, the fixed broadband wireless system will be built as part of a commercial wireless business in which the system capacity represents the inventory of the business – the product that the wireless business has for sale. How the link capacity is packaged is important to how desirable it will be to system customers. A system that very efficiently provides a set of highly reliable slow speed 10-kbps data channels does not provide the operator with a product inventory that is competitive or marketable. From this perspective, the starting point in calculating system capacity is market research, which in turn defines the required service and application inventory. The engineering details of the system are then determined so that the required inventory is provided in the most cost-effective way.

The required inventory can be quite varied with different customers requiring different data rates, service quality, and availability. The terms of what the wireless operator is providing (and what the customer is paying for) are sometimes embodied in service level agreements (SLAs). Lower data rates, quality or availability carry a lower price. Technological flexibility that allows the service to be packaged in a wide variety of ways is another aspect of the commercial capacity that should be considered in network design.

Commercial capacity can be directly tied to business revenue potential and ultimately to whether there is a viable business case for building the system. If the business case is not initially attractive, modifications to the engineering elements that affect capacity and cost as listed in Section 8.9.2 can be made in an effort to create a system that is commercially feasible.

8.10 CONCLUSION

Fixed broadband wireless systems fall into two broad categories. The first category includes individual links used to connect two distinct points. The only access question is whether an interference-free frequency can be found to carry the service. In the second, more general, case, there will be multiple users contending for use of the same spectrum. These users will be both cooperative (within the control of the system operator), and uncooperative as in the case of a competing system in the same geographical area or in a neighboring area.

For the case of cooperative multiple users in the same system, this chapter has described a variety of approaches to harmoniously allowing many users to successfully access the spectrum simultaneously. The techniques described include FDMA, TDMA, CDMA, SDMA, OFDMA, and CSMA. Each is intended to exploit particular dimensions of the spectrum space, but more importantly, modern systems have evolved in which two or more of these approaches are used in combination to achieve system capacity that is well suited to the network requirements.

SDMA is particularly universal as it can be added in combination with any of the other described methods without the use of additional bandwidth or time. The cost associated with SDMA is hardware complexity; a cost which continually decreases with more efficient and powerful signal processing chips. SDMA also requires more physical space for the adaptive antenna elements. Nevertheless, of the multiple-access techniques described here, SDMA offers the greatest future potential for increasing system capacity when compared to the potential of the other technologies.

8.11 REFERENCES

[1] H.R. Anderson, "Digicast: toward a more efficient use of the radio spectrum resource," *Intelligent Machines Journal*, no. 10, page 12, June 1979.

[2] Federal Communications Commission (FCC) MM Docket 97-217, "Methods for predicting interference from response stations transmitters and to response station hubs and for supplying data on response station systems," April 21, 2000. (Appendix D).

[3] Telecommunications Industry Association (TIA/EIA), "Wireless Communications – Performance in Noise, Interference-Limited Situations – Recommended Methods for Technology-Independent Modeling, Simulation and Verification." TSB-88-1, June 1999, and Addendum 1, January, 2002.

[4] IEEE Computer Society Working Group 802.16, "IEEE Recommended Practice for Local and Metropolitan Area Networks – Coexistence of Fixed Broadband Wireless Access Systems." *Institute of Electrical and Electronic Engineers*, September, 2001.

[5] S.C. Yang. *CDMA RF System Engineering*. Boston: Artech House, 1998. Chapter 7.

[6] S. Verdu, "Minimum probability of error for asynchronous Gaussian multi-access channels," *IEEE Transactions on Information Theory*, vol. 32, pp. 85–96, January, 1986.

[7] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 35, pp. 123–136, January, 1989.

[8] M. Haardt, A. Klein, R. Koehn, S. Oestreich, M. Purat, V. Sommer and T. Ulrich, "The TD-CDMA based UTRA TDD mode," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1375–1385, August, 2000.

[9] Tim Wilkinson. Private communication. June, 2002.
[10] 3GPP web site. *http://www.3gpp.org*.
[11] H. Holma, S. Heikkinen, O.A. Lehtinen, and A. Toskala, "Interference considerations for the time division duplex mode of the UMTS terrestrial radio access," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1386–1393, August, 2000.
[12] R. van Nee and R. Prasad. *OFDM for Wireless Multimedia Communications.* Boston: Artech House, 1998. Chapter 8.
[13] IEEE P802.16.a/D4-2002. "Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems – Medium Access Control Modifications and Additional Physical Layer Specifications for 2–11 GHz, IEEE Standard for Local and Metropolitan Area Networks: Coexistence of Fixed Broadband Wireless Access Systems," *Institute of Electrical and Electronic Engineers*, May 27, 2002.
[14] V.K. Garg and J.E. Wilkes. *Wireless and Personal Communications Systems.* Upper Saddle River, N.J.: Prentice Hall PTR. 1996.
[15] B.P. Lathi. *Modern Digital and Analog Communication Systems*. New York: Holt, Reinhart and Winston. 1983.