ADVANCES IN
MULTIUSER DETECTION
ADVANCES IN MULTIUSER DETECTION

Edited by

Michael L. Honig
Northwestern University
Preface xv
Contributors xvii

1 Overview of Multiuser Detection 1

Michael L. Honig

1.1 Introduction / 1
  1.1.1 Applications / 2
  1.1.2 Mobile Cellular Challenges / 4
  1.1.3 Chapter Outline / 5
1.2 Matrix Channel Model / 6
1.3 Optimal Multiuser Detection / 8
  1.3.1 Maximum Likelihood (ML) / 8
  1.3.2 Optimal (Maximum a Posteriori) Detection / 9
  1.3.3 Sphere Decoder / 10
1.4 Linear Detectors / 12
  1.4.1 Comparison with Optimal Detection / 13
  1.4.2 Properties of Linear Multiuser Detection / 15
1.5 Reduced-Rank Estimation / 16
  1.5.1 Subspaces from the Matched Filter / 17
  1.5.2 Eigen-Space Methods / 18
    1.5.2.1 Principal Components (PC) / 18
    1.5.2.2 Generalized Side-lobe Canceller (GSC) / 18
    1.5.2.3 Cross-Spectral Method / 19
    1.5.2.4 Comparison / 19
  1.5.3 Krylov Subspace Methods / 20
    1.5.3.1 Multi-Stage Wiener Filter (MSWF) / 20
    1.5.3.2 Rank-Recursive (Conjugate Gradient) Algorithm / 22
    1.5.3.3 Performance / 22
    1.5.3.4 Adaptive Rank Selection / 24
  1.5.4 Performance Comparison / 25
2 Iterative Techniques

*Alex Grant and Lars K. Rasmussen*

2.1 Introduction / 47
   2.1.1 System Model / 48
   2.1.2 Multiuser Detectors / 50
      2.1.2.1 Optimal Multiuser Detectors / 51
      2.1.2.2 Decorrelator Detector / 52
      2.1.2.3 Linear Minimum Mean-Squared Error Detectors / 52
      2.1.2.4 Per-User Linear Minimum Mean-Squared Error Detectors / 53
      2.1.2.5 Per-User Approximate Nonlinear MMSE Detector / 54
   2.2 Iterative Joint Detection for Uncoded Data / 56
      2.2.1 Interference Cancellation / 56
         2.2.1.1 Schedules for Iterative Cancellation / 58
         2.2.1.2 Implementation via Residual Error Update / 59
         2.2.1.3 Tentative Decision Functions / 63
      2.2.2 Linear Methods / 64
         2.2.2.1 Solutions to Linear Systems / 66
         2.2.2.2 Direct Solution / 66
         2.2.2.3 Series Expansions / 67
         2.2.2.4 Iterative Solution Methods / 70
         2.2.2.5 Jacobi Iteration / 72
         2.2.2.6 Gauss-Seidel Iteration / 74
         2.2.2.7 Descent Algorithms / 75
2.2.3 Non-Linear Methods / 80
   2.2.3.1 Constrained Optimization / 80
   2.2.3.2 Clipped Soft Decision / 84
   2.2.3.3 Hyperbolic Tangent / 86
   2.2.3.4 Hard Decision / 86
2.2.4 Numerical Results / 88
   2.2.4.1 Parallel Cancellation / 88
   2.2.4.2 Serial Cancellation / 89
   2.2.4.3 Gradient Methods / 93
2.3 Iterative Joint Decoding for Coded Data / 95
   2.3.1 Joint Optimal Multiuser and Separate Single-User Decoders / 96
   2.3.2 The Canonical Iterative Joint Multiuser Decoder / 97
   2.3.3 Linear Detection in Iterative Joint Multiuser Decoding / 100
   2.3.4 Parallel Interference Cancellation / 102
   2.3.5 Per-User LMMSE Filters with Priors / 103
   2.3.6 Transfer Function Convergence Analysis / 105
   2.3.7 Numerical Examples / 107
      2.3.7.1 Separate Multiuser Detection and Single-User Decoding / 107
      2.3.7.2 Single-User Matched Filter Parallel Interference Cancellation / 107
      2.3.7.3 Per-User LMMSE Filtering / 111
      2.3.7.4 Comparison of Single-User Matched-Filter PIC and LMMSE Decoders / 115
2.4 Concluding Remarks / 118
References / 119

3 Blind Multiuser Detection in Fading Channels
   Daryl Reynolds, H. Vincent Poor, and Xiaodong Wang

   3.1 Introduction / 127
   3.2 Signal Models and Blind Multiuser Detectors for Fading Channels / 129
      3.2.1 Asynchronous Multi-Antenna Multipath CDMA / 129
      3.2.2 Synchronous Multipath CDMA / 134
      3.2.3 Synchronous Multi-Antenna CDMA / 136
      3.2.4 Remarks / 137
   3.3 Performance of Blind Multiuser Detectors / 138
      3.3.1 Complex Gaussian Distribution / 138
      3.3.2 Performance of Blind Multiuser Detectors with Known Channels / 139
      3.3.3 Performance of Blind Multiuser Detector with Blind Channel Estimation / 142
3.3.4 Numerical Results / 143
3.3.5 Adaptive Implementation / 144
3.3.6 Algorithm Summary / 146

3.4 Bayesian Multiuser Detection for Long-Code CDMA / 148
3.4.1 System Descriptions / 148
  3.4.1.1 Signal and Channel Model / 148
  3.4.1.2 Noise Model / 149
  3.4.1.3 Blind Bayesian Multiuser Detection / 150
  3.4.1.4 The Gibbs Sampler / 151
3.4.2 Bayesian MCMC Multiuser Detectors / 152
  3.4.2.1 White Gaussian Noise / 152
  3.4.2.2 Colored Gaussian Noise / 155
3.4.3 Simulation Examples / 157

3.5 Multiuser Detection for Long-Code CDMA in Fast-Fading Channels / 161
3.5.1 Channel Model and Sequential EM Algorithm / 161
3.5.2 Sequential Blind Multiuser Detector / 163
3.5.3 Simulation Results / 163

3.6 Transmitter-Based Multiuser Precoding for Fading Channels / 165
3.6.1 Basic Approach and Adaptation / 166
  3.6.1.1 Uplink Signal Model and Blind Channel Estimation / 166
  3.6.1.2 Downlink Signal Model and Matched Filter Detection / 166
  3.6.1.3 Transmitter Precoding for a Synchronous Multipath Downlink / 167
  3.6.1.4 Adaptive Implementation / 169
  3.6.1.5 Algorithm Summary / 170
3.6.2 Precoding with Multiple Transmit Antennas / 171
  3.6.2.1 Downlink Signal Model / 171
  3.6.2.2 Precoder Design for Orthogonal Spreading Codes / 171
  3.6.2.3 Precoder Design for Non-Orthogonal Spreading Codes / 172
3.6.3 Precoding for Multipath ISI Channels / 174
  3.6.3.1 Prerake-Diversity Combining / 174
  3.6.3.2 Precoder Design / 175
3.6.4 Performance Analyses / 178
  3.6.4.1 Performance of Transmitter Precoding with Blind Channel Estimation / 178
  3.6.4.2 Performance and Achievable Diversity for Multi-Antenna Precoding / 180
4 Performance with Random Signatures

Matthew J. M. Peacock, Iain B. Collings, and Michael L. Honig

4.1 Random Signatures and Large System Analysis / 189

4.2 System Models / 192
   4.2.1 Uplink CDMA Without Multipath / 193
   4.2.2 Downlink CDMA / 194
   4.2.3 Multi-Cell Downlink or Multi-Signature Uplink / 196
   4.2.4 Model Limitations / 197

4.3 Large System Limit / 198
   4.3.1 SINR of Linear Filters / 198

4.4 Random Matrix Terminology / 201
   4.4.1 Eigen-Value Distributions / 201
   4.4.2 Stieltjes Transform / 201
   4.4.3 Examples / 202
   4.4.4 Asymptotic Equivalence / 203

4.5 Incremental Matrix Expansion / 204

4.6 Analysis of Downlink Model / 206
   4.6.1 MMSE Receiver and SINR / 206
   4.6.2 Large-System SINR / 207
   4.6.3 Two Important Preliminary Results / 208
      4.6.3.1 Rotational Invariance of SINR / 208
      4.6.3.2 Covariance Matrix Expansion Along Transmit Dimensions / 209
   4.6.4 Large System SINR / 210
   4.6.5 Numerical Example / 213

4.7 Spectral Efficiency / 215
   4.7.1 Sum Capacity / 215
   4.7.2 Capacity Regions / 219

4.8 Adaptive Linear Receivers / 221
   4.8.1 ALS Receiver / 221
   4.8.2 ALS Convergence: Numerical Example / 223
   4.8.3 Large System Limit / 224
   4.8.4 Analysis and Results / 225
      4.8.4.1 ALS Transient Behavior / 226
      4.8.4.2 Steady-State SINR / 228
   4.8.5 Numerical Examples / 228
   4.8.6 Optimization of Training Overhead / 230

4.9 Other Models and Extensions / 236
Appendix: Proof Sketch of Theorem 1 / 238
Appendix: Free Probability Transforms / 241
4.B.1 Free Probability Transforms / 242
4.B.2 Sums of Unitarily Invariant Matrices / 243
4.B.3 Products of Unitarily Invariant Matrices / 245
References / 246

5 Generic Multiuser Detection and Statistical Physics 251

Dongning Guo and Toshiyuki Tanaka

5.1 Introduction / 251
   5.1.1 Generic Multiuser Detection / 251
   5.1.2 Single-User Characterization of Multiuser Systems / 252
   5.1.3 On the Replica Method / 254
   5.1.4 Statistical Inference Using Practical Algorithms / 254
   5.1.5 Statistical Physics and Related Problems / 255

5.2 Generic Multiuser Detection / 256
   5.2.1 CDMA/MIMO Channel Model / 256
   5.2.2 Generic Posterior Mean Estimation / 256
   5.2.3 Specific Detectors as Posterior Mean Estimators / 259
      5.2.3.1 Linear Detectors / 260
      5.2.3.2 Optimal Detectors / 260
      5.2.3.3 Interference Cancelers / 260

5.3 Main Results: Single-User Characterization / 261
   5.3.1 Is the Decision Statistic Gaussian? / 261
   5.3.2 The Decoupling Principle: Individually Optimal Detection / 262
   5.3.3 Decoupling Principle: Generic Multiuser Detection / 269
      5.3.3.1 A Companion Channel / 269
      5.3.3.2 Main Results / 271
   5.3.4 Justification of Results: Sparse Spreading / 272
   5.3.5 Well-Known Detectors as Special Cases / 273
      5.3.5.1 Linear Detectors / 273
      5.3.5.2 Optimal Detectors / 275

5.4 The Replica Analysis of Generic Multiuser Detection / 276
   5.4.1 The Replica Method / 276
      5.4.1.1 Spectral Efficiency and Detection Performance / 276
      5.4.1.2 The Replica Method / 277
      5.4.1.3 A Simple Example / 278
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4.2</td>
<td>Free Energy</td>
<td>281</td>
</tr>
<tr>
<td>5.4.2.1</td>
<td>Large Deviations and Saddle Point</td>
<td>283</td>
</tr>
<tr>
<td>5.4.2.2</td>
<td>Replica Symmetry Solution</td>
<td>285</td>
</tr>
<tr>
<td>5.4.2.3</td>
<td>Single-User Channel Interpretation</td>
<td>286</td>
</tr>
<tr>
<td>5.4.2.4</td>
<td>Spectral Efficiency and Multiuser Efficiency</td>
<td>288</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Joint Moments</td>
<td>289</td>
</tr>
<tr>
<td>5.5</td>
<td>Further Discussion</td>
<td>291</td>
</tr>
<tr>
<td>5.5.1</td>
<td>On Replica Symmetry</td>
<td>291</td>
</tr>
<tr>
<td>5.5.2</td>
<td>On Metastable Solutions</td>
<td>292</td>
</tr>
<tr>
<td>5.6</td>
<td>Statistical Physics and the Replica Method</td>
<td>294</td>
</tr>
<tr>
<td>5.6.1</td>
<td>A Note on Statistical Physics</td>
<td>294</td>
</tr>
<tr>
<td>5.6.2</td>
<td>Multiuser Communications and Statistical Physics</td>
<td>296</td>
</tr>
<tr>
<td>5.6.2.1</td>
<td>Equivalence of Multiuser Systems and Spin Glasses</td>
<td>296</td>
</tr>
<tr>
<td>5.7</td>
<td>Interference Cancellation</td>
<td>297</td>
</tr>
<tr>
<td>5.7.1</td>
<td>Conventional Parallel Interference Cancellation</td>
<td>297</td>
</tr>
<tr>
<td>5.7.2</td>
<td>Belief Propagation</td>
<td>298</td>
</tr>
<tr>
<td>5.7.2.1</td>
<td>Application of Belief Propagation to Multiuser Detection</td>
<td>298</td>
</tr>
<tr>
<td>5.7.2.2</td>
<td>Conventional Parallel Interference Cancellation as Approximate BP</td>
<td>300</td>
</tr>
<tr>
<td>5.7.2.3</td>
<td>BP-Based Parallel Interference Cancellation Algorithm</td>
<td>301</td>
</tr>
<tr>
<td>5.8</td>
<td>Concluding Remarks</td>
<td>303</td>
</tr>
<tr>
<td>5.9</td>
<td>Acknowledgments</td>
<td>304</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>304</td>
</tr>
</tbody>
</table>

# 6 Joint Detection for Multi-Antenna Channels


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>311</td>
</tr>
<tr>
<td>6.2</td>
<td>Wireless Channels: The Multi-Antenna Realm</td>
<td>312</td>
</tr>
<tr>
<td>6.3</td>
<td>Definitions and Preliminaries</td>
<td>314</td>
</tr>
<tr>
<td>6.4</td>
<td>Multi-Antenna Capacity: Ergodic Regime</td>
<td>315</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Input Optimization and Capacity-Achieving Transceiver Architectures</td>
<td>316</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Random Matrix Theory</td>
<td>320</td>
</tr>
<tr>
<td>6.4.2.1</td>
<td>Eigen-Value Distributions</td>
<td>320</td>
</tr>
<tr>
<td>6.4.2.2</td>
<td>Transforms</td>
<td>320</td>
</tr>
</tbody>
</table>
6.4.3 Canonical Model (IID Channel) / 321
   6.4.3.1 Separable Correlation Model / 323

6.5 Multi-Antenna Capacity: Non-Ergodic Regime / 327

6.6 Receiver Architectures and Performance / 330
   6.6.1 Linear Receivers / 330
      6.6.1.1 Zero-Forcing Receiver / 331
      6.6.1.2 Minimum Mean-Square Error Receiver / 337

6.7 Multiuser Multi-Antenna Systems / 345
   6.7.1 Same-Cell Interference and Cooperation / 346
      6.7.1.1 Downlink: Precoding / 347
      6.7.1.2 Uplink: Interference Cancellation / 349
   6.7.2 Other-Cell Interference and Cooperation / 349
      6.7.2.1 Joint Encoding / 350
      6.7.2.2 Base Station Cooperative Scheduling / 350

6.8 Diversity-Multiplexing Tradeoffs and Spatial Adaptation / 351
   6.8.1 Diversity-Multiplexing Tradeoff / 352
   6.8.2 Mode Adaptation: Switching Between Diversity and
      Multiplexing / 353

6.9 Conclusions / 355

References / 355

7 Interference Avoidance for CDMA Systems
   Dimitrie C. Popescu, Sennur Ulukus, Christopher Rose, and Roy Yates

7.1 Introduction / 365
7.2 Interference Avoidance Basics / 367
   7.2.1 Greedy Interference Avoidance: The Eigen-Algorithm / 370
   7.2.2 MMSE Interference Avoidance / 372
   7.2.3 Other Algorithms for Interference Avoidance / 376

7.3 Interference Avoidance over Time-Invariant Channels / 377
   7.3.1 Interference Avoidance with Diagonal Channel Matrices / 379
   7.3.2 Interference Avoidance with General Channel Matrices / 380

7.4 Interference Avoidance in Fading Channels / 384
   7.4.1 Iterative Power and Sequence Optimization in Fading / 388

7.5 Interference Avoidance in Asynchronous Systems / 389
   7.5.1 Interference Avoidance for User Capacity Maximization / 390
7.5.2 Interference Avoidance for Sum Capacity Maximization / 396
7.5.3 TSAC Reduction: Iterative Algorithms / 399

7.6 Feedback Requirements for Interference Avoidance / 401
7.6.1 Codeword Tracking for Interference Avoidance / 401
7.6.2 Reduced-Rank Signatures / 402

7.7 Recent Results on Interference Avoidance / 403
7.7.1 Interference Avoidance and Power Control / 403
7.7.2 Adaptive Interference Avoidance Algorithms / 405

7.8 Summary and Conclusions / 410
References / 411

8 Capacity-Approaching Multiuser Communications Over Multiple Input/Multiple Output Broadcast Channels 417

Uri Erez and Stephan ten Brink

8.1 Introduction / 417
8.2 Many-to-One Multiple Access versus One-to-Many Scalar Broadcast Channels / 418
8.3 Alternative Approach: Dirty Paper Coding / 420
8.3.1 The Dirty Paper Coding Result / 420
8.3.2 DPC vs. SSD Approach for a Coded Interference Signal / 421
8.3.3 Scalar Broadcast Using the DPC Approach / 421

8.4 A Simple $2 \times 2$ Example / 423
8.5 General Gaussian MIMO Broadcast Channels / 428
8.5.1 Vector Dirty Paper Coding: Reduction to Scalar Case / 428
8.5.2 DPC Rate Region / 430

8.6 Coding with Side Information at the Transmitter / 431
8.6.1 A Naive Attempt / 432
8.6.2 Scalar Quantization: Tomlinson-Harashima Precoding / 432
8.6.2.1 Dither Signal / 434
8.6.2.2 Losses of Tomlinson-Harashima Precoding / 434
8.6.2.3 MMSE Scaling / 436
8.6.2.4 One-Dimensional Soft-Symbol Metric / 437

8.6.3 Vector Quantization: Sign-Bit Shaping / 440
8.6.3.1 Lattices / 440
8.6.3.2 Shaping Gain / 440
8.6.3.3 Communication Using Lattices / 441
8.6.3.4 Lattice Precoding at the Transmitter / 442
8.6.3.5 High-Dimensional Lattices from Linear Codes / 443
8.6.3.6 Sign-Bit Shaping / 447
8.6.3.7 Coset Decoding at the Receiver / 448
8.6.3.8 Mutual Information Limits / 450
8.6.4 The Role of Channel Knowledge / 451
  8.6.4.1 Single User vs. Multiuser MIMO / 451
  8.6.4.2 Obtaining Channel Knowledge / 452
8.7 Summary / 452
References / 453
The proliferation of telecommunications systems and services over the past couple of decades has been accompanied by numerous advances in physical layer communications and associated signal processing techniques. Many of these services have in fact been enabled by the increase in spectral efficiency provided by improved modulation, coding, and reception capability. Furthermore, the evolution of communications networks continues to stimulate efforts to push the performance and reliability of these networks to their fundamental limits.

For the most part, many of the recent advances in signal processing methods for communications have been motivated by the evolution of mobile cellular systems, i.e., from first-generation analog systems, introduced in the 1980s, to next (fourth)-generation systems currently being designed. Additional motivation has been provided by other wireless systems and standards, such as wireless local and metropolitan area networks, and also the desire to provide broadband services over existing copper subscriber lines in the telephone network. Although wired channels, such as subscriber lines, do not experience the time variations in received signal strength associated with mobile cellular channels, other challenges remain, such as overcoming frequency-selectivity and efficient spectrum sharing among multiple users with different channel characteristics.

This book reviews recent advances in multiuser detection, which generally refers to methods for detecting digital data associated with multiple interfering signals. These advances comprise some of the signal processing techniques just mentioned, and have been an active area of research and development over the past couple of decades. In the title of this book, “recent” generally refers to advances made over the past ten years, i.e., since the publication of the first book on multiuser detection [87]. Except for the first chapter, which gives a general introduction and overview, each of the eight chapters is contributed by a different set of authors, and is meant to be a self-contained discussion of a particular topic. Namely, Chapter 2 discusses iterative techniques for combined multiuser detection and decoding of error control codes; Chapter 3 discusses multiuser detection in the presence of linear channel impairments, such as multipath and intersymbol interference; Chapters 4 and 5 present performance analysis of multiuser detection methods with random signatures and channels; Chapter 6 discusses the application of joint detection methods to Multi-Input/Multi-Output (MIMO) channels, corresponding to wireless links with multiple antennas at the transmitter...
and/or receiver; Chapter 7 discusses interference avoidance methods at the transmitter (i.e., through the choice of signatures assuming spread spectrum signaling); and Chapter 8 discusses transmitter precoding methods for the MIMO downlink (broadcast channel). A more detailed overview of the chapters is given at the end of Chapter 1.

These topics represent a sampling of major advances that have been made in multiuser detection over the past ten years. Because this research area has been quite active, comprehensive coverage of recent progress would be quite difficult. This book therefore serves as an entry point for exploring ongoing research in multiuser detection and for learning about existing unsolved problems and issues. The intended audience is therefore graduate students in communications, as well as practicing engineers and researchers who are familiar with digital communications at the level of [87] and [60], and wish to gain a deeper understanding of multiuser detection techniques. This area continues to progress and it is our hope that these contributions will stimulate further advances.

MICHAEL HONIG

Evanston, Illinois
July 2009
CONTRIBUTORS

Jeffrey G. Andrews, Department of Electrical and Computer Engineering, University of Texas, Austin, Texas
Iain B. Collings, Wireless Technologies Lab, CSIRO ICT Centre, Sydney, Australia
Uri Erez, Department of Electrical Engineering-Systems, Tel Aviv University, Tel Aviv-Yafo, Israel
Alex Grant, Institute for Telecommunications Research, University of South Australia, Mawson Lakes, South Australia
Dongning Guo, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, Illinois
Robert W. Heath, Jr., Department of Electrical and Computer Engineering, University of Texas, Austin, Texas
Michael L. Honig, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, Illinois
Matthew R. McKay, Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong
Matthew J. M. Peacock, Credit-Suisse Bank, New York, New York
H. Vincent Poor, Department of Electrical Engineering, Princeton University, Princeton, New Jersey
Dimitrie C. Popescu, Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia
Lars K. Rasmussen, KTH, Royal Institute of Technology, School of Electrical Engineering, Stockholm, Sweden
Daryl Reynolds, Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, West Virginia
Christopher Rose, Wireless Information Network Laboratory (WINLAB), Rutgers University, North Brunswick, New Jersey
Toshiyuki Tanaka, Department of Systems Science, Kyoto University, Kyoto, Japan
Stephan ten Brink, Wionics Technologies, Irvine, California

Antonia Tulino, Dip. Di Ing. Elettronica e delle Telecomunicazioni, Università degli Studi di Napoli, Naples, Italy

Sennur Ulukus, Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland

Xiaodong Wang, Electrical Engineering Department, Columbia University, New York, New York

Roy Yates, Wireless Information Network Laboratory (WINLAB), Rutgers University, North Brunswick, New Jersey
OVERVIEW OF MULTIUSER DETECTION

Michael L. Honig

. . . But what is the use of counterpoint when, if played, one imagines that four different orchestras are playing at the same time four different tunes in four different keys and measures? A veritable nightmare!

—Music critic Arthur Bird, writing about R. Strauss’s tone poem *Ein Heldenleben*, May 1, 1899 [69, p. 186]

1.1 INTRODUCTION

One of the key challenges in designing multiuser communications systems is mitigating interference. This challenge is apparent for modern wireless networks, such as mobile cellular, and wireless local and metropolitan area networks, where achieving high spectral efficiencies requires aggressive frequency reuse. Interference also limits the performance of many wired channels, such as the digital subscriber line (DSL). Although each DSL is typically associated with a single user, capacitive coupling between pairs of DSLs in close physical proximity causes cross-talk interference, which degrades performance (e.g., see [72]).

Interference encompasses self-interference, due to reflections of the same signal, in addition to multi-user-interference associated with other signals sharing the same bandwidth. Self-interference arises from multipath in a wireless channel, a bridge tap in a Digital Subscriber Line (DSL), and bandwidth constraints, which cause...
intersymbol interference. Throughout this book, interference typically refers to signals from other users (or more generally, data streams) associated with the same system. Clearly, techniques for reducing, or mitigating interference lead directly to improved performance, either in terms of reduced error rate, increased data rate, or number of users that can be served.

Effective techniques for mitigating interference must depend, of course, on what is known at the transmitter and receiver. For example, in a multiple-access channel the receivers for all transmitters are co-located. Hence instead of detecting the transmitted bits from a particular user in isolation, treating the other signals as background noise, the receiver can jointly detect all of the transmitted information bits, or symbols. In that case, the structure of the multiple access interference can be used to aid the detection of the desired symbol. For example, it may be possible to use estimates of interfering symbols to cancel, or at least reduce, the level of the interfering signals. In contrast, if the receivers are not co-located, then interference cancellation, which relies on accurate interference estimates, may not be practical.

Multiuser detection refers to the scenario in which a single receiver jointly detects multiple simultaneous transmissions. Examples include the uplink of a single cell in a cellular system, and a group of twisted-pair copper subscriber lines that terminate at the same central switching office. More generally, multiuser detection techniques apply to the joint detection of different signals transmitted over any multi-input/multi-output (MIMO) channel. In addition to the preceding examples, others include channels in which multiple transmitted information streams are multiplexed over multiple transmit antennas. In that scenario, the multiple “users” refer to the multiple information streams, even though the transmitted signal may originate from a single user.

Closely related to multiuser detection is interference suppression. The key distinction is that a multiuser detector attempts to retrieve multiple (i.e., at least two) transmitted signals, or information streams, whereas interference suppression implies that the receiver is interested in only one signal among the received superposition of transmitted signals. For example, this is typically the case for the downlink of a cellular system, in which a mobile wishes to demodulate a single transmitted information stream in the presence of interfering signals from the associated base station and from nearby base stations. Of course, in general a receiver may wish to demodulate a subset of two or more signals from among a larger mix of signals. In that case, the receiver jointly detects the subset of desired signals while suppressing the interfering signals. An example of this is the uplink of a cellular system, in which the receiver is interested in demodulating the transmitted signals from users within the cell in the presence of interference from other cells.

1.1.1 Applications

Much of the work on multiuser detection and interference suppression over the past couple of decades has been motivated by the commercial success of Code-Division Multiple Access (CDMA) in mobile cellular systems. Namely, second-generation CDMA cellular systems were introduced in the early 1990s, and CDMA is currently used in third-generation cellular systems. The performance of CDMA is generally
limited by multi-user interference. In particular, the performance of the uplink, which refers to the multiple access channel from users to base station, is sensitive to received power variations across users. The classic example of this is the near–far problem, in which a user close to the base station causes excessive interference to a user far from the base station. Commercial CDMA systems generally use closed-loop power control to minimize received power variations (both across time to mitigate fading and across users). It was recognized early on, however, that the sensitivity to interference power is not inherent to CDMA, but rather is a property of the conventional matched filter detector. That motivated studies of multiuser detection techniques for uplink CDMA, starting with the work of Verdú [86].

Prior to the introduction of CDMA for mobile cellular systems, some specific multiuser detectors had been derived for linear MIMO channels. See, for example, [80], which discusses the Maximum Likelihood (ML) detector, and [1,45,46,63], which derive linear and decision feedback detectors.

In recent years, the primary motivation for multiuser detection has shifted from CDMA to other applications (in particular, links with multiple antennas). There are several reasons for this. First, as shown in the initial studies on the achievable rates for wireless links with multiple antennas [19,77], substantial gains in spectral efficiency and reliability can be achieved by adding antennas to the transmitter and/or receiver of a single-user wireless channel. That stimulated an enormous amount of activity on coding and reception techniques for MIMO channels, so that nearly all evolving wireless systems and standards include provisions for multiple transmit antennas. In those scenarios, multiuser detection techniques are useful for mitigating interference among the different transmit antennas.

The second reason for the shift away from CDMA applications is that current designs for next-generation cellular and wide-area wireless networks are based primarily on Orthogonal Frequency Division Multiplexing (OFDM) and Orthogonal Frequency Division Multiple Access (OFDMA). That is due in part to the more substantial role played by data services, as opposed to voice, in evolving wireless networks. Namely, current CDMA cellular systems rely on “interference averaging” and power control to achieve robustness with respect to variations in the active user set and associated channels. That requires a relatively large number of low-rate (e.g., voice) users. In contrast, data traffic associated with internet services is bursty, and depends on the rates provided. Higher rates enable shorter bursts, but typically generate more interference due to higher transmit power. Hence without further coordination among users, the interference seen at a base station or mobile is likely to vary substantially as high-rate users enter and leave, degrading performance.

1Although [86] stimulated much of the subsequent work on multiuser detection, the structure of the optimal detector for CDMA had been previously derived in [67].

2Although that work predates the development of mobile cellular and DSL, similar types of “multi-terminal” applications are mentioned, including multi-cable and diversity channels. The MIMO models in [1,45] were originally motivated by dual polarization radio transmission with crosstalk.

3In principle, fluctuations in interference in a CDMA system can be mitigated through the application of multiuser detection, although the challenges subsequently discussed still apply.
To mitigate interference associated with bursty users and high data rates, transmissions can be scheduled over different time slots (scheduling intervals). Because data is delay-tolerant, that also allows the scheduler to select users with favorable channel conditions, thereby exploiting multiuser diversity \cite{3,5,49}. Hence current (third-generation) cellular data systems typically rely on scheduling, as opposed to spreading (as in CDMA), to reduce interference. Emerging systems based on OFDMA provide further flexibility in allocating both time and frequency resources among requests to mitigate interference and exploit multiuser diversity. Interference mitigation techniques are potentially useful for OFDMA, although the interference originates from other cells (or sectors), so that associated channel conditions may be more difficult to estimate.

1.1.2 Mobile Cellular Challenges

Cellular systems pose several challenges that prevent a straightforward application of most multiuser detection techniques. This is especially true for uplink asynchronous CDMA with full spreading, or frequency reuse. The main difficulty is that for interference averaging, it is desirable to extend spreading sequences over many symbols, so that interference from a particular user is averaged over many (randomly chosen) signatures. In that way, the performance is not limited by the possibility of choosing a particular set of signatures with undesirable correlation properties. These “long” signatures, however, greatly complicate the implementation of a multiuser detector, which exploits properties of the particular set of assigned signatures for each symbol. That is, the structure of the multiuser detector must change from symbol to symbol, which may require excessive computation. Long signatures also prevent the application of standard adaptive filtering methods (used to equalize single-user channels), which require linear modulation with short signatures (i.e., the signature for a particular user is repeated from symbol to symbol).\footnote{At a basic level, one might argue that interference averaging as a design objective for cellular systems runs contrary to a design based on multiuser detection. Namely, interference averaging is associated with a large number of weak interferers, whereas multiuser detection becomes most attractive when there are a small number of strong interferers.}

Another challenge for multiuser detection posed by mobile cellular systems is fading. The multiuser detector must track or adapt to channel variations in time and frequency caused by mobility. That includes channel variations associated with the interferers as well as the desired user. This may be feasible when the channel variations are slowly varying (e.g., over a few hundred symbols), or when the number of interferers is relatively small, so that the number of channel parameters to track is manageable. However, in a CDMA cellular system there may be a large number of mobile users, several of which experience rapid fading. The inability to track all of these channel variations can significantly compromise performance. Furthermore, the complexity of the multiuser detector generally increases with the system size (number of users and processing gain). For this reason, application of multiuser detection to uplink CDMA has been mostly limited to relatively simple interference cancellation techniques.
One way to address the preceding challenges is to redesign the CDMA system with multiuser detection in mind. In addition to using short signatures, that generally means reducing the number of users, or transmitted information streams, which are jointly detected at the base station, and also slowing down the channel variations (fade rate). Reducing the number of users for joint detection can be accomplished by sub-dividing the channel resources in time or frequency. For example, the Time-Division Duplex (TDD) version of UMTS$^5$ is based on a combination of Time-Division Multiple Access (TDMA) and CDMA [30]. The uplink channel is divided into 10 msec frames with 15 time slots, and multiple users can be assigned to a particular time slot through the use of direct-sequence spread spectrum signaling (i.e., direct-sequence CDMA). For example, if the original CDMA system supports 75 users in a cell, then introducing 15 time slots means that the users can be divided into 15 groups, each containing five users, which are assigned to the different time slots. That reduces the required spreading (processing gain) by approximately a factor of 15, and eases the burden on a multiuser detector. Furthermore, the introduction of time slots increases the symbol rate, and hence decreases the fade rate (rate of channel variations normalized by the symbol rate). That simplifies the associated channel estimation.

As previously discussed, the desire to integrate voice services with different data services having variable Quality of Service (e.g., delay) requirements has motivated the trend towards OFDM and OFDMA. The channel is then divided into both time and frequency slots, and each slot is designated for a single user. Hence in general the transmitted signals are not spread. That obviates the need for multiuser detection on the uplink, although interference from other cells still limits performance. Hence interference suppression techniques may still be beneficial. Nevertheless, multiuser detection and interference suppression for OFDM systems are primarily focused on mitigating interference among multiple antennas, which creates MIMO channels across the sub-carriers.

### 1.1.3 Chapter Outline

In the next section we discuss the linear (matrix) channel model, which is the basis for the multiuser detection methods discussed in this book. We then give a brief overview of a few different multiuser detectors along with performance comparisons for a simple version of this channel model in which the mixing matrix has independent, identically distributed (i.i.d.) elements. This discussion is meant to provide some background and motivation for the topics discussed in subsequent chapters. We also discuss and compare some additional multiuser detection techniques. A comprehensive treatment of some of the detectors presented here (e.g., optimal and linear) is given in the book by Verdú [87].

---

$^5$Universal Mobile Telecommunications System is an air interface standard for third generation cellular systems. The standard was approved in 1998, and consists of two modes corresponding to Frequency-Division Duplex (FDD) and TDD operation [30,57].
1.2 MATRIX CHANNEL MODEL

In its simplest and most general form the matrix channel model is given by:

\[ y(i) = M(i)b(i) + n(i) \]  \hspace{1cm} (1.1)

where \( y \) is an \( N \times 1 \) vector of received samples, \( b \) is a \( K \times 1 \) vector of transmitted symbols, \( M \) is an \( N \times K \) mixing matrix, and \( n \) is an \( N \times 1 \) vector of additive noise samples. Also, \( i \) is the discrete time index. Hence this discrete-time model includes the combination of any filtering (analog or digital) at the transmitters and receivers. Assuming quadrature modulation, all of the variables in (1.1) are complex valued.

For synchronous CDMA the \( k \)th entry of \( b(i) \), denoted \( b_k(i) \), is the \( i \)th symbol transmitted by user \( k \), and the \( k \)th column of \( M \) is the signature for user \( k \). In that case the elements of \( M \) are typically chosen from a discrete set, e.g., \( \pm 1 \) for real codes, or \( \pm 1 \pm j \) for complex signatures. In practice, the signature elements are often chosen randomly, although it is also desirable if the signatures (columns of \( M \)) are orthogonal. To represent different powers across the users, we can write \( M = SA \), where \( S \) is the signature matrix, and \( A \) is a diagonal matrix of amplitudes.

For asynchronous CDMA, the model (1.1) still applies, but with different interpretations for \( b(i) \) and \( M \). Namely, assuming each signature spans a single symbol interval, we can write:

\[ y(i) = M_1b(i - 1) + M_2b(i) + M_3b(i + 1) + n = Mb(i) + n \] \hspace{1cm} (1.2)

where the \( k \)th columns of \( M_1, M_2, \) and \( M_3 \) are the corresponding segments of signatures (padded with zeros) associated with \( b_k(i - 1), b_k(i), \) and \( b_k(i + 1) \), respectively, \( b(i) \) contains \( b(i - 1), b(i), \) and \( b(i + 1) \) stacked on top of each other, and \( M \) is the corresponding block-diagonal matrix.

For a single-user link with \( K \) transmit antennas and \( N \) receive antennas, if \( b(i) \) contains \( K \) symbols (i.e., is \( K \times 1 \)), then \( b_k \) in (1.1) is transmitted by the \( k \)th antenna, and the \((n,k)\)th entry of \( M \) is the complex channel gain from the \( k \)th transmit antenna to the \( n \)th receive antenna. More generally, the number of transmitted symbols \( K' \) [dimension of \( b(i) \)] may be less than the number of transmit antennas \( K \), in which case an additional \( K \times K' \) precoding matrix \( V \) is needed to map the symbols to antennas. That is, \( M = HV \), where \( H \) is the \( N \times K \) channel matrix. The channel gains are often modeled as i.i.d. random variables (typically complex Gaussian corresponding to flat Rayleigh fading). Other statistical models for the matrix \( M \), corresponding to correlated fading across antennas, are discussed in Chapter 6.

The interference channel can also be modeled by (1.1), where \( y \) corresponds to the signal at a particular receiver, which estimates a subset of \( b \). The remaining symbols are presumably transmitted to other receivers at different locations, and are treated as interference. That applies, for example, to the uplink where the estimated symbols are transmitted by users within the cell and the interfering symbols are transmitted by users in other cells.
The model (1.1) also applies to a single-input/single-output channel with intersymbol interference (ISI). Namely, suppose a single transmitter transmits the sequence of symbols \( \{b(i)\} \) through a linear, dispersive channel with impulse response:

\[
h(-m), h(-m + 1), \ldots, h(0), h(1), \ldots, h(m),
\]

where the length of the impulse response is assumed to be \( 2m + 1 \), so that the output:

\[
y(i) = \sum_{k=-m}^{m} h(k)b(i-k)
\]

(1.3)

We can represent this in the form (1.1) by defining the vector of \( N = 2n + 1 \) received samples, corresponding to time \( i \), as:

\[
y(i) = [y(i+n), \ldots, y(i), \ldots, y(i-n)]^T.
\]

(1.4)

where \([\cdot]^T\) denotes transpose, and the vector of \( K = 2(n + m) + 1 \) transmitted symbols as \( b(i) = [b(i+n+m), b(i+n+m-1), \ldots, b(i-n-m)]^T \). The channel matrix \( M \) is then the \( N \times K \) Toeplitz matrix:

\[
M = \begin{bmatrix}
h(-m) & \ldots & h(m) & 0 & \ldots & 0 \\
0 & h(-m) & \ldots & h(m) & \ldots & 0 \\
& & \ddots & \ddots & & \\
0 & \ldots & 0 & h(-m) & \ldots & h(m)
\end{bmatrix}.
\]

(1.5)

In practice, the receiver would presumably use \( y(i) \) to detect the subset of symbols \( b(i+n-m), \ldots, b(i-n+m) \), which forms the reduced symbol vector \( \bar{b}(i) \). That is, the receiver estimates only the subset of transmitted symbols, which correspond to columns of \( M \) containing the entire impulse response. (That maximizes the received energy associated with each detected symbol.) Assuming that symbols are transmitted continuously, the trailing \( 2m \) symbols of \( \bar{b} \), namely, \( b(i-n+m-1), \ldots, b(i-n-m) \), therefore experience intersymbol interference from the preceding vector \( \bar{b}(i-1) \), and the leading \( 2m \) symbols \( b(i+n+m), \ldots, b(i+n-m+1) \) interfere with \( \bar{b}(i+1) \).

To model an OFDM system, the channel matrix \( M \) is first converted to a circulant matrix by setting the first \( m \) symbols of \( \bar{b} \) equal to the trailing \( m \) symbols of \( \bar{b}(i) \), and the trailing \( m \) symbols of \( \bar{b} \) equal to the first \( m \) symbols of \( \bar{b}(i) \) [i.e., \( b(i+n+m) = b(i-n+2m-1), \ldots, b(i+n+1) = b(i-n+m) \) and \( b(i-n-1) = b(i+n-m), \ldots, b(i-n-m) = b(i+n-2m+1) \)]. If we rewrite (1.1) in terms of \( \bar{b}(i) \), i.e.:

\[
y(i) = \bar{M}\bar{b}(i) + n(i),
\]

(1.6)
then $\mathbf{M}$ is circulant, i.e., its rows are cyclic shifts of the first row:

$$h(0), \ldots, h(m), 0, \ldots, 0, h(-m), \ldots, h(-1).$$

The matrix $\mathbf{M}$ can be diagonalized by pre-multiplying $\mathbf{b}$ by the inverse DFT matrix $\mathbf{W}$, and post-multiplying $y(i)$ by the DFT matrix $\mathbf{W}^\dagger$. Hence the conclusion is that (1.1) models a single-user OFDM channel, where the mixing matrix $\mathbf{M}$ can be assumed to be diagonal. The diagonal entries are the complex channel gains across sub-channels.

Other interpretations of the model (1.1) that pertain to uplink and downlink CDMA with dispersive channels are discussed in Chapter 4. In those scenarios, the mixing matrix $\mathbf{M}$ is the product of a Toeplitz, circulant, or diagonal channel matrix, depending on the implementation, and a signature matrix. It becomes apparent that other network configurations with multiple users, multiple antennas, dispersive channels, and with or without spreading can be modeled by (1.1), where $\mathbf{M}$ and $\mathbf{b}(i)$ take on the appropriate forms. For purposes of the following overview, we will focus on the preceding interpretations of the model (1.1), which apply to synchronous CDMA, multi-antenna links (MIMO channels), and dispersive Single-Input/Single-Output (SISO) channels.

### 1.3 OPTIMAL MULTIUSER DETECTION

We now provide a brief overview of multiuser detection techniques that have been proposed for the matrix model (1.1), starting in this section with optimal detectors. The intent is to provide some background and points of reference for the advances and performance results presented in subsequent chapters. We also discuss some additional related techniques.

#### 1.3.1 Maximum Likelihood (ML)

Referring to (1.1), the ML detector chooses the vector of estimated symbols as:

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}} \Pr(y \text{ received} | \mathbf{b} \text{ transmitted})$$

(1.7)

where the dependence on $i$ has been dropped for convenience. If the additive noise $\mathbf{n}$ is Gaussian, then this is equivalent to selecting:

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}} \|y - \mathbf{M}\mathbf{b}\|$$

(1.8)

where the norm is the regular Euclidean norm and the elements of $\mathbf{b}$ are constrained to be constellation points.

---

6Alternatively, the transmitted packet can be rearranged so that the first $2m$ symbols are the same as the trailing $2m$ symbols. The first $2m$ symbols are then the “cyclic prefix” [24, Ch. 12].
The introduction of the ML detector for CDMA in [86] essentially launched the field of multiuser detection. The main reason for this is that it was shown that the performance of the ML detector is insensitive to power variations among the users, unlike the matched filter detector. Hence this suggested that improving the CDMA receiver could relax the requirements on closed-loop power control, which was one of the most complex aspects of CDMA system design.

Although the ML detector can nearly eliminate the degradation in performance due to multiuser interference for low to moderate loads $K/N$ (users per degree of freedom), it has two main drawbacks: complexity and required side information. Because $b$ takes on discrete values, computing the estimate $\hat{b}$ is an optimization over a discrete set, which is known to be computationally difficult. Specifically, the computation associated with known algorithms for determining the minimum in (1.1), assuming that $b$ and $y$ can be chosen arbitrarily, grows exponentially with $K$ (the size of the vector $b$), corresponding to the exponential growth in the size of the set over which the minimization is taken. This is not a major issue if the dimension of $b$ is small (say, $K<10$ with binary signaling), so that ML detection may be appropriate for single multi-antenna links, or for a TDMA system in which there are relatively few co-channel users per time slot. However, the ML search complexity clearly becomes impractical for a CDMA system with much more than ten users per sector.

The second issue with the ML detector is that (1.1) implicitly assumes that $M$ is known. Again, this is probably not a major issue for multi-antenna links, which experience slow fading, so that the channel can be accurately estimated. Even so, for some wireless applications, ML multiuser detection may not be as attractive as other simpler decision feedback techniques, to be discussed. The reason is that when combined with error control coding, the combined ML detector for the error control code concatenated with the channel becomes prohibitively complex. In that case, it is desirable to produce soft estimates of the transmitted symbols, i.e., with reliability information. This is further discussed later in this chapter and in Chapter 2.

1.3.2 Optimal (Maximum a Posteriori) Detection

The Maximum a Posteriori (MAP) detector selects:

$$\hat{b} = \arg \max_b \Pr(b \text{ transmitted} \mid y \text{ received}), \quad (1.9)$$

which minimizes the probability of error. This is the same as the ML estimate if the symbols are equally likely. However, when combined with error control coding and iterative soft decoding, the decoder can pass reliability information to the multiuser detector in the form of the $a$ priori distribution, or likelihood ratio for each transmitted symbol. Hence, in that scenario the MAP estimate generally differs from the ML estimate. Furthermore, the MAP detector itself computes soft estimates of each symbol (e.g., likelihood ratios), although the final (hard) estimates are obtained from the soft estimates by thresholding.
If the receiver detects a subset of the vector $b$, then the MAP estimate maximizes the corresponding marginal distribution [e.g., $Pr(b_k | y)$ for a particular user $k$]. In general, this differs from the estimate in (1.9) and requires less computation.

For asynchronous CDMA, the multiuser MAP detector can be implemented using the standard forward-backward algorithm (e.g., see [87, Sec. 4.2]). The MAP detector suffers from the same drawbacks as the ML detector, namely, the complexity grows exponentially with the size of $b$, and it requires knowledge of $M$. However, in some applications where the system size is relatively small, the complexity may be manageable.

### 1.3.3 Sphere Decoder

Because of the high complexity of the optimal detector, it is generally desirable to incur some performance loss in order to simplify the receiver. That trade off motivates the linear and decision feedback detectors to be discussed. Other reduced-complexity detectors have been proposed that essentially approximate the optimal detector (e.g., see [51,70,75]), or rely on specific properties of the signatures (e.g., [64,66]).

Although the worst-case complexity of the ML search increases exponentially with $K$ (e.g., see [87, Ch. 4]), the average complexity, taking into account the model (1.1), may be much less. An example of a search algorithm with relatively low average complexity, assuming the Gaussian noise model, is the sphere decoder, discussed in [10,33].

To describe the sphere decoder, we first observe that if each component of $b$ is selected from a rectangular (e.g., QAM) constellation, then each $b$ corresponds to a point in a rectangular lattice. The ML estimate in (1.8) can then be computed by performing an exhaustive search over that lattice. To reduce the search complexity, the sphere decoder restricts $b$ to lie in a hypersphere of radius $r$ centered at $y$. If the hypersphere contains at least one lattice point, then this restricted search still gives the ML estimate.

Given $r$, an algorithm for finding all points in a lattice within the hypersphere defined by:

$$\|b - \hat{b}\|^2 M M (b - \hat{b}) \leq r^2,$$

where $\hat{b}$ is the center of the sphere, was presented in [16]. Application of this sphere decoding algorithm to decoding a lattice code was presented in [11], and was subsequently proposed for MIMO channels in [10,33,90]. (See also [76], which considers a somewhat more general search constraint.) For convenience, we assume that all variables are real-valued (hence Hermitian transpose $\dagger$ will be replaced by transpose $T$). The extension to rectangular QAM constellations is obtained by stacking the real and imaginary components of each complex vector (see [10]). The extension to circular (PSK) constellations is discussed in [33].

---

7“Worst-case” refers to the scenario in which $y$ and $b$ can be chosen to maximize the search time, given a particular search algorithm.