To my family
Acknowledgements

I wish to express my sincere gratitude to my advisor, Professor H. Vincent Poor, for his guidance and support throughout the course of my doctoral study. I would also like to thank the professors of the Departments of Electrical Engineering, Operation Research and Financial Engineering, and Mathematics, whose courses have provided the solid background and foundation for my thesis research. Among them, I give special thanks to Professor Stuart Schwartz and Professor S-Y. Kung for serving on my committee, reading my work and giving me valuable comments, and to Professor Sergio Verdú and Professor Hisashi Kobayashi for their inspiring teaching and advice.

I want to take this opportunity to thank Dr. Reinaldo Valenzuela and Dr. Justin Chuang for summer internships at Bell Labs, Lucent Technologies, and at AT&T Labs-Research, respectively. I also would like to thank Dr. Laurence Mailaender and Dr. Andreas Molisch for their mentoring and collaborations during these internships. These industry experiences have greatly benefited my research and future career.

Furthermore, I am indebted to my parents for their advice, support and love. I would also like to acknowledge the contribution of all my former teachers for their cultivation and encouragement. Finally, I thank my lovely baby girl for the hope and joy brought by her, and my wife for her wordless support and selfless love.
Abstract

Wireless cellular communications has been one of the fastest growing fields of technology in the world. As opposed to its wireline counterparts, wireless communications poses some unique challenges including multipath fading and co-channel interference. Diversity techniques are overwhelmingly used in wireless communication systems to enhance capacity, coverage and quality, among which space diversity, i.e., diversity realized in space with antenna arrays, is favored because it does not impose a penalty in terms of scarce spectrum resources. Under the framework of wireless cellular communications with antenna arrays, both signal processing and information theoretic aspects are studied in this dissertation.

The signal processing techniques investigated are, among others, space-time processing, multiuser detection, and turbo decoding. All of these techniques exhibit near-Shannon-limit performance with reasonable complexities in many cases, and are very promising for next-generation communications. Specifically, various transmit diversity and downlink beamforming techniques with power control are examined and compared for wireless cellular communications with transmit arrays, and a range of iterative space-time multiuser detection techniques are explored with receive arrays. Further, turbo space-time multiuser detection techniques are employed for wireless cellular multiple-input multiple-output (MIMO) communications, i.e., with antenna arrays on both transmit and receive ends. Then, for multicell MIMO systems where co-channel interference is the dominating detrimental factor, various multiuser receivers are proposed to dramatically improve the system performance.
Spectral efficiency of MIMO systems operating in multicell frequency-flat fading environments is also studied. The following detectors are analyzed: a single-cell detector, the joint optimum detector, a group linear minimum-mean-square-error (MMSE) detector, a group MMSE successive cancellation detector, and an adaptive multiuser detector. Large-system asymptotic (non-random) expressions for their spectral efficiencies are developed. Some analytical and numerical results are derived based on these expressions to gain insight into the behavior of multicell MIMO systems.

Even though wireless cellular communications constitutes the main part of this dissertation, an application of some of the methods developed to wireline communications is also considered. In particular, the turbo multiuser detection techniques are applied to digital subscriber line (DSL) wireline communications to effectively combat crosstalk, with the influence of impulse noise taken into consideration.


### Contents

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
</tbody>
</table>

#### 1 INTRODUCTION

1.1 OVERVIEW ................................................................. 1

1.2 DISSERTATION OUTLINE AND CONTRIBUTIONS ............................................. 7

#### 2 TRANSMIT ARRAYS: DOWNLINK BEAMFORMING WITH POWER CONTROL ................................................................. 11

2.1 INTRODUCTION .......................................................................................... 11

2.2 SYSTEM MODEL ......................................................................................... 15

2.2.1 Multipath Channel ................................................................................... 15

2.2.2 FDD Framework ....................................................................................... 19

2.2.3 Cellular System ....................................................................................... 20

2.3 POWER CONTROL/ALLOCATION ALGORITHMS ................................................. 21

2.3.1 Perron-Frobenius Theorem and its Applications ...................................... 22

2.3.2 General Form of Power Control Problem ............................................... 23

2.4 ARRAY SIGNAL PROCESSING ....................................................................... 25

2.4.1 Transmit Diversity ................................................................................... 25

2.4.2 Sectorization ........................................................................................... 28

2.4.3 Beamforming Techniques ......................................................................... 29

2.4.3.1 Beam Steering ....................................................................................... 31

2.4.3.2 Maximum SNR ..................................................................................... 32

2.4.3.3 Maximum SIR/SINR ............................................................................ 32

2.4.4 Joint Power Control and Maximum SINR Beamforming ............................. 33

2.5 NUMERICAL RESULTS ............................................................................ 36

2.5.1 Circuit-Switched System ......................................................................... 37

2.5.2 Packet-Switched System ......................................................................... 46

2.6 SUMMARY ................................................................................................... 53
### 3 RECEIVE ARRAYS: ITERATIVE SPACE-TIME MULTIUSER DETECTION

3.1 INTRODUCTION ........................................................................................................ 54
3.2 SPACE-TIME SIGNAL MODEL ............................................................................ 56
3.3 BATCH ITERATIVE METHODS ................................................................................ 60
  3.3.1 *Iterative Linear ST MUD* ................................................................. 60
  3.3.2 *Iterative Nonlinear ST MUD* .............................................................. 63
    3.3.2.1 Cholesky Iterative Decorrelating Decision-Feedback ST MUD .... 64
    3.3.2.2 Multistage Interference Cancelling ST MUD ............................. 68
  3.3.3 *EM-based Iterative ST MUD with a New Structure* ......................... 68
    3.3.3.1 EM and SAGE Algorithm with Application to ST MUD .......... 69
    3.3.3.2 SAGE Iterative ST MUD with a New Structure ....................... 73
  3.3.4 Numerical Results .................................................................................... 77
3.4 SAMPLE-BY-SAMPLE ADAPTIVE METHODS .................................................... 85
  3.4.1 *Data-Aided ST MUD* ........................................................................... 85
    3.4.1.1 Decentralized Adaptive MMSE ST MUD ............................... 86
    3.4.1.2 Centralized Adaptive Decision-Feedback ST MUD ............... 90
    3.4.1.3 Numerical Results ................................................................. 91
  3.4.2 *Blind ST MUD* .................................................................................... 95
    3.4.2.1 LCMV Blind ST MUD and its GSC Implementation .......... 95
    3.4.2.2 MIN-MAX Channel Parameter Estimation ......................... 99
    3.4.2.3 Robustified Blind ST MUD .................................................. 100
    3.4.2.4 Numerical Results ................................................................. 101
3.5 SUMMARY ....................................................................................................... 104

### 4 MIMO SYSTEMS: TURBO SPACE-TIME MULTIUSER DETECTION

4.1 INTRODUCTION ................................................................................................. 106
4.2 PROBLEM FORMULATION ............................................................................ 108
  4.2.1 *MIMO System Model* ........................................................................ 108
  4.2.2 *Cellular System Model* ........................................................................ 110
4.3 TURBO SPACE-TIME MULTIUSER DETECTION FOR INTRACELL COMMUNICATIONS
   .......................................................................................................................... 113
  4.3.1 Receiver Structures and Diversity ............................................................ 113
  4.3.2 Turbo-BLAST Detection ........................................................................ 116
4.4 MULTIUSER DETECTION TO COMBAT INTERCELL INTERFERENCE ......... 124
6.3.1 Maximum Likelihood Multiuser Detection ........................................................ 188
6.3.2 Interference Cancellation Multiuser Detection ................................................. 188
6.3.3 Robust Multiuser Detection with Impulse Noise............................................. 191
6.3.4 Importance Sampling Techniques for Intensive Simulations......................... 193
6.4 Turbo Multiuser Detection for Coded DSL ....................................................... 196
  6.4.1 Turbo Decoding for Coded DMT System....................................................... 199
6.5 Numerical Results ............................................................................................... 201
  6.5.1 Robust Multiuser Detection with Impulse Noise........................................... 201
  6.5.2 Turbo Multiuser Detection........................................................................... 205
6.6 Summary ........................................................................................................... 216

7 Conclusions and Perspectives ........................................................................... 218

Bibliography .......................................................................................................... 223
List of Tables

Table 2.1 Number of users supported in a cell with 5% outage ........................................ 39
Table 2.2 Mean SINR (50% CDF) of a typical mobile user (dB) ............................................. 48
Table 2.3 Peak SINR (90% CDF) of a typical mobile user (dB) .............................................. 48
List of Figures

Fig. 2.1 Cellular simulation model ................................................................. 21
Fig. 2.2 Sector antenna radiation pattern ...................................................... 28
Fig. 2.3 Performance comparison of various transmission techniques with M = 2 antennas per sector (6 antennas per cell) — circuit-switched system ...................... 38
Fig. 2.4 Performance comparison of various transmission techniques with M = 4 antennas per sector (12 antennas per cell) — circuit-switched system .................... 38
Fig. 2.5 Performance comparison of various transmission techniques with M = 8 antennas per sector (24 antennas per cell) — circuit-switched system .............. 39
Fig. 2.6 Performance of transmit diversity with 2, 4 and 8 antennas per sector .......... 41
Fig. 2.7 Performance of transmit diversity with sectorization with 2, 4 and 8 antennas per sector ........................................................................................................ 41
Fig. 2.8 Performance of max SNR beamforming without feedback with 2, 4 and 8 antennas per sector........................................................................................................ 42
Fig. 2.9 Performance of max SNR beamforming with feedback with 2, 4 and 8 antennas per sector........................................................................................................ 42
Fig. 2.10 Performance of beam steering with 2, 4 and 8 antennas per sector .......... 43
Fig. 2.11 Performance of max SNR beamforming with 2 antennas per sector: with and without feedback channel information...................................................... 44
Fig. 2.12 Performance of max SNR beamforming with 4 antennas per sector: with and without feedback channel information...................................................... 45
Fig. 2.13 Performance of max SNR beamforming with 8 antennas per sector: with and without feedback channel information...................................................... 45
Fig. 2.14 Performance comparison of various transmission techniques with M = 4 antennas and 2 Active users — packet-switched system................................. 46
Fig. 2.15 Performance comparison of various transmission techniques with M = 8 antennas and 2 active users — packet-switched system................................. 47
Fig. 2.16 Performance comparison of various transmission techniques with M = 8 antennas and 4 active users — packet-switched system................................. 47
Fig. 2.17 Performance of max SNR beamforming with 4 antennas and 2 active users: with and without feedback channel information........................................ 50
Fig. 2.18 Performance of max SNR beamforming with 8 antennas and 2 active users: with and without feedback channel information........................................ 50
Fig. 2.19 Performance of max SNR beamforming with 8 antennas and 4 active users: with and without feedback channel information........................................ 51
Fig. 2.20 Performance of max SINR beamforming with 4 antennas and 2 active users: with and without feedback channel information ................................................ 51
Fig. 2.21 Performance of max SINR beamforming with 8 antennas and 2 active users: with and without feedback channel information ................................................ 52
Fig. 2.22 Performance of max SINR beamforming with 8 antennas and 4 active users: with and without feedback channel information ................................................ 52
Fig. 3.1 A conventional space-time multiuser receiver structure ........................................ 59
Fig. 3.2 Cholesky iterative decorrelating decision-feedback ST MUD ................................ 66
Fig. 3.3 A new space-time multiuser receiver structure ..................................................... 73
Fig. 3.4 Performance comparison of BER versus SNR for five space-time multiuser receivers ............................................................................................................... 81
Fig. 3.5 Performance comparison of decision-feedback ST MUD and Cholesky iterative ST MUD .............................................................................................................. 82
Fig. 3.6 Performance comparison of SAGE ST MUD with traditional and new ST structure .................................................................................................................. 83
Fig. 3.7 Performance comparison of convergence behavior of multistage interference cancelling ST MUD and EM-based iterative ST MUD ........................................... 84
Fig. 3.8 Structure of an adaptive MMSE space-time multiuser detector ................................ 87
Fig. 3.9 Structure of an adaptive centralized decision-feedback space-time multiuser detector ................................................................................................................... 90
Fig. 3.10 Convergence of the decentralized adaptive MMSE space-time multiuser detector ...................................................................................................................... 92
Fig. 3.11 Bit error rate of the decentralized adaptive MMSE space-time multiuser detector in the steady state ........................................................................................................ 93
Fig. 3.12 Comparison of steady state output SINR of the two adaptive receivers .............. 94
Fig. 3.13 Comparison of steady state BER of the two adaptive receivers ............................. 94
Fig. 3.14 GSC implementation of LCMV blind adaptive space-time multiuser detector (pth antenna FIR for desired user k) ........................................................................ 97
Fig. 3.15 Convergence of the LCMV blind adaptive space-time multiuser detector ......... 102
Fig. 3.16 Precision of min/max channel parameter estimation ............................................ 102
Fig. 3.17 Performance comparison of LCMV blind adaptive ST MUD with exact and estimated channel parameters ................................................................. 103
Fig. 3.18 Performance comparison of non-robust and robust LCMV blind adaptive ST MUD in the situation of signature waveform mismatch ................................. 103
Fig. 4.1 Cellular system with one tier of interferers in the downlink case ...................... 111
Fig. 4.2 Structure of coded V-BLAST ................................................................................ 115
Fig. 4.3 Structure of Turbo-BLAST ................................................................. 116
Fig. 4.4 CDF of SNR experienced by a mobile ................................................. 128
Fig. 4.5 CDF of SIR experienced by a mobile ..................................................... 128
Fig. 4.6 CDF of the ratio between the power sum of the two strongest interferers and the power sum of the remaining interferers experienced by a mobile ....................... 129
Fig. 4.7 CDF of the ratio between the power of the strongest interferer and the power sum of the remaining interferers experienced by a mobile ........................................ 129
Fig. 4.8 Performance comparison of various MIMO receivers when two equal-power interferers dominate ................................................................................... 131
Fig. 4.9 Performance comparison of various versions of group IC MUD when two equal-power interferers dominate ........................................................................ 133
Fig. 4.10 Performance comparison of various MIMO receivers when one interferer dominates ........................................................................................................ 133
Fig. 4.11 Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 1dB .................................................. 135
Fig. 4.12 Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 3dB .................................................. 136
Fig. 4.13 Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 5dB .................................................. 136
Fig. 4.14 Downlink capacity of interference-limited MIMO when one interferer dominates ........................................................................................................ 137
Fig. 4.15 Downlink capacity of interference-limited MIMO when two interferers dominate ........................................................................................................ 138
Fig. 4.16 Comparison of theoretical and simulated results of the capacity of interference-limited MIMO Systems with linear MMSE front end ....................................... 138
Fig. 4.17 CDF of block error rate for different receivers experienced by a mobile in Rayleigh fading ..................................................................................................... 140
Fig. 4.18 CDF of block error rate for different receivers experienced by a mobile in Ricean fading ..................................................................................................... 142
Fig. 5.1 Study of interference-limited behavior of various multicell MIMO spectral efficiencies ........................................................................................................ 172
Fig. 5.2 Spectral efficiency comparison of group linear MMSE decoder, group MMSE successive cancellation decoder, and adaptive decoder ........................................ 175
Fig. 5.3 Multicell MIMO capacity for different SNRs and SIRs .......................... 176
Fig. 6.1 VDSL DMT System Configuration: (a) Transmitter; (b) Channel; (c) Receiver ......................................................................................................................... 184
Fig. 6.2 Huber penalty function and its derivative for the Gaussian mixture model 
$\varepsilon = 0.1, \kappa = 100, \sigma^2 = 1, k = 1.14$ .................................................................... 192
Fig. 6.3 Turbo structure for iterative demodulation and decoding ........................................ 197
Fig. 6.4 Bit error rate (BER) versus signal-to-noise ratio (SNR) for different detectors (x-mark: SUD, circle: IC-MUD, diamond: ML-MUD, dashed: single user lower bound) ................................................................................................................ 203
Fig. 6.5 Bit error rate (BER) versus signal-to-noise ratio (SNR) for different detectors (x-mark: SUD, circle: IC-MUD, plus: IC-MUD-R, diamond: ML-MUD, star: ML-MUD-R, dashed: single user lower bound). .................................................................................... 204
Fig. 6.6 Bit allocation for DMT subchannels ........................................................................ 207
Fig. 6.7 Performance of the ML-MAP turbo multiuser receiver ........................................ 208
Fig. 6.8 Performance of the IC-MAP turbo multiuser receiver ........................................ 208
Fig. 6.9 Performance of the ML-SOVA turbo multiuser receiver ..................................... 209
Fig. 6.10 Performance of the IC-SOVA turbo multiuser receiver .................................... 209
Fig. 6.11 Performance comparison of various DMT VDSL receivers ............................ 211
Fig. 6.12 Performance of the iterative DMT receiver with Gray coding ....................... 213
Fig. 6.13 Performance of the iterative DMT receiver with natural coding ..................... 213
Fig. 6.14 Gray coding for 4-PAM .................................................................................. 214
Fig. 6.15 Natural coding for 4-PAM ............................................................................. 214
Fig. 6.16 Performance of the iterative DMT receiver with impulse noise .................... 214
Fig. 6.17 Performance of the iterative DMT receiver with erasure decoding with impulse noise ................................................................................................................................. 216
Chapter 1

Introduction

1.1 Overview

The wireless era began around 1897 when Guglielmo Marconi first established a radio link to provide continuous contact with ships sailing the English Channel. Since then, mobile systems have developed and spread considerably. The use of wireless systems increased rapidly in the 1950’s and 60’s, when the number of users largely exceeded the small number of channels available. This trend showed a clear need for larger capacity as well as better roaming flexibility. Bell Labs addressed these issues in 1970’s by proposing a new conceptual idea called the cellular concept – the concept of breaking a coverage zone into small cells, each of which reuses portions of the spectrum to increase the user capacity [86]. From then on, wireless cellular communications has been one of the fastest growing fields of technology in the world.

The first-generation (1G) cellular and cordless telephone networks, which were based on analog technology with FM modulation, have been successfully deployed throughout the world since the early and mid-1980s. But the remarkable growth of the market was such that a second-generation (2G) of wireless systems was quickly needed. Due to the well known advantages of digital transmission, especially in the use of digital
modulation, speech coding, and spectrally efficient multiple access schemes such as frequency-division multiple-access (FDMA), time-division multiple-access (TDMA), and code-division multiple-access (CDMA), the 2G digital cellular radio showed a significant improvement in both system capacity and quality of service over the 1G systems. As the economy and technologies keep growing, the demand for a third-generation (3G) of wireless communication systems becomes increasingly urgent. These systems are evolving from the mature 2G networks, with the aim of providing universal access, global roaming, and multimedia high-speed high-quality wireless communications.

Wireless communications, as opposed to their wireline counterpart, pose some unique challenges, among them are:

(1) The radio propagation environment and the mobility of users give rise to multipath echoes and signal fading in time, frequency, and space.

(2) Cellular communications and some multiple access schemes (e.g., CDMA) bring about strong co-channel interference.

(3) The limited allocated spectrum results in a limit on capacity.

(4) The limited battery life at the mobile end poses restrictions on the hardware design and software signal processing complexity.

Diversity techniques are overwhelmingly used in wireless communication systems to combat the multipath fading. The principle of diversity is to use a number of transmission paths, all carrying the same signal but having independent fading statistics. Proper combination of the signals from all the paths yields a result with greatly reduced severity
of fading and thus improved reliability of transmission. Diversity can be realized in time, frequency, or code, as well as in space with antenna arrays.

Space diversity is favored for mobile radio use because it does not require a penalty in terms of scarce spectrum resources. It can be applied in several different ways. If we have antenna arrays at the transmit end only, it is called multiple-input single-output (MISO) diversity. If the total transmit power is constrained, the capacity gain of a MISO system diminishes with an increase in the number of antennas. On the other hand, if we have antenna arrays at the receive end only, it is called single-input multiple-output (SIMO) diversity. The capacity of a SIMO system increases logarithmically with the number of antennas. The use of antenna arrays at both ends of a communication link is the key idea of multiple-input multiple-output (MIMO) systems, whose capacity grows linearly with the number of antennas. Furthermore, depending on the type of fading that is to be mitigated, space diversity techniques can be divided in two groups: microscopic diversity and macroscopic diversity. The use of small closely spaced multi-element antenna arrays to combat the small scale fading (multipath) belongs to the microscopic diversity category. The use of base station antennas widely separated in space to combat the large scale fading (shadowing) is a form of macroscopic diversity.

In a fading environment, the antenna elements should be separated sufficiently far apart to experience uncorrelated fading and get the diversity gain. In urban area, the required spacing is half a wavelength at the mobile and ten times a wavelength at the base station. In the indoor environment, however, half wavelength spacing is enough for both ends of a MIMO link. Independent of the fading environment and in additional to the diversity gain, multiple antennas can provide antenna gain due to the potential coherent
combining of the transmitted and/or received signals and the underlying uncorrelated
noise. This technique is often called beamforming, where the signals are modeled as
planar wavefronts impinging on/transmitting from an antenna array with a certain
direction of arrival (DOA)/ direction of departure (DOD). In this approach, the signal
structure induced by multiple antennas, i.e., the spatial signature, can be exploited for co-
channel interference suppression. Note that in a quasi-static Rayleigh fading
environment, which is exemplified by the indoor wireless systems, we can assume both
that the fading coefficients in different antennas are independent, so as to get diversity
gain, and that the receiver can learn these coefficients so as to exploit the spatial signature
to produce the antenna gain [42].

Space-time processing, with its ability to improve the capacity, coverage and quality
over time-alone processing in wireless networks by reducing co-channel interference
while enhancing diversity and array gain, draws increasing interest of both academia and
industry recently [78]. There are various ways to exploit the spatial dimension, which can
roughly be divided into two categories. One is space-time coding, aiming at approaching
the wireless system’s capacity in adverse environments with the design of good space-
time codes [104]. The current study of space-time coding is focused on the transmit end
with only one or two antenna elements at the receiver. As we mentioned before, the
capacity of such system is limited and does not increase without bound as the number of
transmit antennas increases. Moreover, the decoding complexity of the space-time trellis
codes is rather high. Another approach, the Bell Labs space-time layered architecture
(BLAST) [41], is somewhat complementary. It concentrates on increasing the system
capacity through exploiting a large number of antennas on both ends of a communication
link. Then, instead of endeavoring to approach the system capacity, it is satisfied with achieving a hefty portion of the resulting substantial capacity through signal processing of reasonable complexity.

Actually, the space-time layered architecture falls into the larger category of *space-time multiuser detection*. Multiuser detection (MUD) [119] deals with the detection of data from some or all users when observed in a mutually interfering environment. It exploits the well-defined structure of the multiuser interference, distinct from that of ambient noise, in order to improve the system performance with which channel resources are employed. Multiuser detection can be applied naturally in CDMA systems using nonorthogonal spreading codes. It also can be employed in wireless TDMA or FDMA systems due to the effects of non-ideal channelization or multipath, or to combat co-channel interference from adjacent cells. Multiuser detection techniques include the optimum maximum-likelihood joint detection and various suboptimum linear and non-linear methods. Linear multiuser detection, including decorrelating (zero-forcing) MUD and minimum-mean-square-error (MMSE) MUD, is relatively simple and effective, but its performance is limited in overloaded (more users than degrees of freedom) systems. Non-linear multiuser detection such as decision feedback MUD and multistage interference cancellation MUD, often serves as a favorable tradeoff between performance and complexity. Space-time multiuser detection refers to the application of the multiuser detection techniques above with the aid of both temporal (e.g. CDMA codes) and spatial (spatial signature) structures of the signals to be detected. Note that the BLAST technique is actually a decision feedback space-time multiuser detector.
Error control coding [25] is a common way of approaching the capacity of communication channels and has moved from being a mathematical curiosity to being a fundamental element in the design of modern digital communication systems. Recent trends in coding favor parallel and/or serially concatenated coding and probabilistic, soft-decision, iterative (turbo-style) decoding techniques, which exhibit near-Shannon-limit performance with reasonable complexities in many cases [8], [32]. This promising technique, turbo decoding, will find many applications in emerging communications applications that require moderate error rates and can tolerate a certain amount of decoding delay.

Almost all digital communication systems nowadays have both coding and modulation components. Almost always the desired signals are received in a non-orthogonally multiplexed environment, whether the interference comes from desired or undesired sources. Therefore, multiuser detection is widely applicable in demodulation, although the tradeoff between efficiency and complexity must be taken into consideration. By introducing an interleaver between coding and modulation to form a serially concatenated coding system at the transmitter, and the associated turbo decoding between the multiuser detector and channel decoder at the receiver, the idea of turbo multiuser detection [79] has drawn much attention recently. In this thesis, space-time processing, multiuser detection, and turbo decoding will be jointly studied under the framework of turbo space-time multiuser detection, more of which will be discussed in Chapters 2-5.

Even though wireless cellular communications constitutes the major part of this dissertation, we also consider an application of similar methodologies in wireline
communications. It is expected that emerging communication infrastructures will use high capacity wired media in the metropolitan and wide area environments and employ wireless media in the local area to fulfill the high-quality seamless communications anytime and anywhere. The above-mentioned multiuser detection and turbo decoding techniques are readily carried on to the wireline digital subscriber line (DSL) systems, as discussed in Chapter 6.

1.2 Dissertation Outline and Contributions

This dissertation is organized as follows.

In Chapter 2, wireless cellular communications with transmit antenna arrays is studied. Transmit diversity and various beamforming techniques are studied and compared, in conjunction with power control techniques. No instant downlink channel information is assumed; however, the obtained results are also compared with results assuming ideal feedback. The study is carried out for both the circuit-switched and packet-switched systems, where different conclusions are drawn. The results of this study are reported in [146].

In Chapter 3, space-time multiuser detection for wireless communications with receive arrays is studied. To overcome the computational burden that rises very quickly with increasing numbers of users and receive antennas in asynchronous multipath CDMA channels, efficient implementations of space-time multiuser detection algorithms are considered here. Batch iterative methods assume knowledge of all signals and channels and are suitable for base station processing in cellular systems. They include iterative
linear space-time multiuser detection, Cholesky iterative decorrelating decision-feedback space-time multiuser detection, multistage interference cancelling space-time multiuser detection, and EM-based iterative space-time multiuser detection. Sample-by-sample adaptive methods require knowledge only of the signal and (possibly) channel of a desired user and are specifically suitable for mobile-end processing. Sample-by-sample adaptive methods are also useful for base station processing due to the time varying nature of mobile communications. They include both data aided (with training sequences) and blind methods. For data aided adaptive methods, a decentralized adaptive MMSE space-time multiuser detector and a centralized adaptive decision-feedback space-time multiuser detector are presented. For blind methods, a blind adaptive space-time multiuser receiver based on the linear constrained minimum variance (LCMV) criterion and min-max parameter estimation is developed, which is robustified with norm-constrained techniques in the case of signature waveform mismatch. The results in this chapter have been published in [142], [143], or are to be published in [138].

In Chapter 4, Turbo space-time multiuser detection for wireless cellular MIMO systems is studied, which has come remarkably close to the ultimate capacity limits in Gaussian ambient noise for an isolated cell. Then it is combined with various multiuser detection methods for combating intercell interference. Among various multiuser detection techniques examined, linear MMSE and successive interference cancellation have been shown to be feasible and effective. Based on these two multiuser detection schemes, one of which may outperform the other for different settings, an adaptive detection scheme is developed, which together with a Turbo space-time multiuser detection structure offers substantial performance gain over the well known V-BLAST
techniques with coding in the interference-limited cellular environment. The obtained multiuser capacity is excellent in high to medium signal-to-interference ratio (SIR) scenario. Nonetheless, numerical results also indicate that a further increase in system complexity, using base-station cooperation, could lead to further significant increases of the system capacity. Some of the results in this chapter have been published in [141], or are to be published in [137].

In Chapter 5, the spectral efficiency of wireless cellular MIMO systems operating in multicell frequency-flat fading environments, where co-channel interference is the dominant channel impairment, is studied. The following detectors are analyzed: a single-cell detector, the joint optimum detector, a group linear MMSE detector, a group MMSE successive cancellation detector, and an adaptive multiuser detector. Large-system asymptotic (non-random) expressions for these spectral efficiencies are also explored. Some analytical and numerical results are derived based on the asymptotic multicell MIMO spectral efficiencies to gain insight into the behavior of multicell MIMO systems.

In Chapter 6, turbo multiuser detection for wireline DSL communications is studied. The traditional single-user data detector for such systems merges crosstalk into the background noise, which is assumed to be white and Gaussian. Here we consider the application of multiuser detection and turbo decoding in a discrete multi-tone (DMT) very-high-rate DSL (VDSL) system to combat crosstalk and to obtain substantial coding gain. The effects of impulse noise are also examined, which has been found to greatly impact the performance of multiuser receivers. Two approaches are taken to mitigate the influence of the impulse noise, one is the robust multiuser detection, and the other is an
erasure decoding technique. The results in this chapter have been published in [139], [140], [144], and [145].

Finally, Chapter 7 contains conclusions and perspectives on open problems and future work.
Chapter 2

Transmit Arrays:
Downlink Beamforming with Power Control

2.1 Introduction

In this chapter, wireless cellular communications with transmit antenna arrays is studied. Transmit diversity and various beamforming techniques are explored and compared, in conjunction with power control techniques.

Cellular base stations may make use of an antenna array to achieve diversity gains or antenna gains so as to improve system capacity. In a fading environment, the antenna elements should be separated sufficiently far apart to experience uncorrelated fading and get the diversity gain. Independent of the fading environment and in additional to the diversity gain, multiple antennas can provide antenna gain due to the potential coherent combining of the transmitted and/or received signals and the underlying uncorrelated noise. While considerable progress has been made on the use of receive arrays for the uplink of cellular systems, comparatively little progress has been made for downlink communications, where instead transmit arrays are exploited at the base stations and only one receiver antenna is used at each mobile handset. Since uplinks and downlinks are
used in duplex mode, it is possible to apply the principle of reciprocity, which implies that the channel is identical on both links, as long as both channels use the same frequency and time instant. In time division duplex (TDD) systems the principle of reciprocity can be applied if the dwell (“Ping-Pong”) time is short compared to the channel coherence time. In frequency division duplex (FDD) systems (FDD is adopted in most current cellular systems), the separation between the uplink and downlink carrier frequencies is large enough to reject the reciprocity principle. However, if the frequency separation is not too large, the uplink and downlink will still share many common features, among which are the number of radio paths, their delays and angles, the large-scale path loss and shadowing, and the variance of small-scale fading [78], [82]. Nevertheless, the instantaneous small-scale fading of the two links is uncorrelated, which makes the downlink problem more difficult for FDD systems. The signal received at the base station provides a means for directly estimating the uplink, not the downlink channel. While such information could be available via a feedback channel from the mobile, we will assume that no such channel exists. The fact that the array response is also frequency dependent further complicates the problem. In this study, we focus on array processing techniques to improve the cellular CDMA downlink, which is foreseen to be of crucial importance for the third generation communication systems supporting wireless Internet, video on demand, and multimedia services.

Perhaps the simplest form of spatial processing is open loop transmit diversity, which will be used as the performance baseline in this study [54], [57]. Sectorization, which can be interpreted as fixed beam transmission, has been shown to be an effective way to improve the system capacity [86]. Other array processing techniques belong to the
beamforming category, where the signals are modeled as planar wavefronts impinging on/transmitting from an antenna array with a certain direction of arrival (DOA)/direction of departure (DOD). A simple form of transmit beamforming is beam steering, which assumes the knowledge of the mobile’s position and forms a beam in the direction of line-of-sight. The performance of beam steering degrades in multipath channels with angle spread. A more sophisticated use of the array is to determine the antenna weighting vector that maximizes signal-to-noise ratio (SNR) at the mobiles. Alternatively, one can borrow the idea from uplink receive array processing and come up with a maximum signal-to-interference ratio (SIR) solution for weighting vector design, i.e., maximizing the ratio of the received power of the signal at the desired user and that leaked to the other users. The key element that comes into play of the max SNR or max SIR scheme is the spatial covariance matrix, more details of which will be given later. Note that, compared with its counterpart on uplink processing, there are two differences for the max SIR scheme: 1) the interference term is what this signal contributes to the other users, not that seen at the desired mobile; 2) the power levels of transmitted signals are not available at this stage (it is decided at the power allocation step discussed in the following), so we can not do max signal-to-interference-and-noise ratio (SINR) as uplink processing.

Power control was conceived originally as a mechanism to deal with the near-far problem, but a more general emerging view is that it is a flexible mechanism to provide different quality-of-service to users with heterogeneous requirements [52]. For downlink transmission, power control is also important for energy conservation and interference mitigation. Standard power control algorithms have been reported in [40], [45], [46].
When we perform the above downlink transmission array processing together with power control, we execute it in two steps: 1) an array weighting vector is determined (not needed for transmit diversity) and the SINR is calculated (as functions of transmitted powers) for each mobile receiver; 2) transmitted power is allocated among users so as to minimize the total transmitted power from the base station while keeping the SINR of all links above a certain threshold.

As we mentioned, the downlink communication scenario is different from that of uplink. While in the uplink the weighting vector designs for different users are decoupled, optimal beamforming for the downlink will have to be considered jointly, because the weighting vector for one user will impact the interference received by other users as well as the useful signal power received by the desired user. The idea of joint power control and downlink beamforming was proposed in [87], [88]. The algorithm in [88] was later modified in [121] to give the optimal transmit beamforming vectors. These algorithms require knowledge of the downlink channel. We make modifications so that only information available from the uplink measurements is used. The result of the original algorithm will serve as the baseline for performance comparison of various techniques.

Besides the lack of direct downlink channel information, the limited number of available antennas may also hamper the algorithms. When the number of antennas is small compared to the number of mobiles (as in a circuit-switched system), there are an insufficient number of degrees of freedom to produce simultaneous nulls for each user. However, in packet-switched systems, where users are delay tolerant, the base station can also control the number of simultaneous transmissions. This implies that the performance
tradeoffs between these algorithms depend on the nature of the traffic. In the packet-switched case, we carry out rate control instead of power control, so we assume the base station will transmit at its maximum power. Thus, the max SIR scheme will be replaced by max SINR in the packet-switched system case.

To summarize, the essential question we are addressing is: for a given number of available transmit antennas, should we use transmit diversity, sectorization, simple directional beam steering, max SNR beamforming, max SIR/SINR beamforming, or a joint beamforming and power control scheme? Which choices are best for circuit and packet systems? These questions are addressed in the context of a system that does not have explicit feedback of the downlink channel measurements.

This chapter is organized as follows: in Section 2.2 we describe the multipath channel model and discuss the spatial covariance matrix approximation for the downlink in FDD systems. A standard power control algorithm is addressed in Section 2.3, and in Section 2.4 various transmit array-processing techniques are discussed. Section 2.5 provides the numerical comparison results for circuit-switched and packet-switched systems. Section 2.6 summarizes this chapter.

2.2 System Model

2.2.1 Multipath Channel

We first introduce the model for array processing. The model for transmit diversity will be discussed in the sequel. In our system setting, each mobile user employs a single
antenna, and communicates with a base station having an $M$-element antenna array. The physical channel between the mobile and the base station is assumed to be wide sense stationary with uncorrelated scattering (WSSUS) multipath frequency-selective fading. The uplink received signal vector at the base station is given by

$$\mathbf{x}(t) = \sum_{k=1}^{K} \sqrt{P_k b_k} \sum_{l=1}^{L_k} F_{kl}^U(t) a^U(\theta_{kl}) c_k(t - \tau_{kl}) + \mathbf{n}(t),$$  \hspace{1cm} (2.1)$$

where $P_k$ is the power transmitted by the $k$th user, $b_k$ is the transmitted data for user $k$, and where $F_{kl}^U(t)$ and $\tau_{kl}$ are the complex gain and delay of the $l$th path of the $k$th user, respectively. $c_k(t)$ is the spreading waveform assigned to user $k$. The path complex gain can be modeled as a random process of the form:

$$F_{kl}^U(t) = \frac{C}{d_k^{\eta/2}} \sqrt{S_k} \alpha_{kl}^U(t),$$  \hspace{1cm} (2.2)$$

where $C$ is a constant, $d_k$ is the distance between the mobile and the base station, and $\eta$ is a path loss parameter. $S_k$ denotes log-normal shadowing, which is assumed to be quasi-time-invariant within the period of interest and frequency independent. $\alpha_{kl}^U(t)$ describes the small-scale fading random process, which is frequency dependent and largely uncorrelated for uplink and downlink. $\theta_{kl}$ is the arrival angle of the $l$th path of the $k$th user. In our model, we assume that $\theta_{kl}$, $1 \leq l \leq L_k$, has a Gaussian distribution centered at $\theta_k$, the direction from the line-of-sight of user $k$. $a(\theta)$ is the array response to a wave impinging from an azimuth direction $\theta$. With the assumptions of planar waves and a uniform linear array, the frequency dependent array response is given by
where $d_a$ is the inter-element spacing of the antenna array. So we have

$$
\mathbf{a}(\theta, f) = [1, e^{-j2\pi d_a\frac{f}{c}\sin(\theta)}, \ldots, e^{-j2\pi d_a(M-1)\frac{f}{c}\sin(\theta)}]^T.
$$

(2.3)\\

where

$$
\mathbf{a}^U(\theta_u) = [1, e^{-j2\pi d_a\frac{f}{c}\sin(\theta_u)}, \ldots, e^{-j2\pi d_a(M-1)\frac{f}{c}\sin(\theta_u)}]^T.
$$

(2.4)\\
n(t)$ is an $M$-dimensional complex Gaussian vector with independent and identically distributed (i.i.d.) components of zero mean and variance $\sigma^2$. When considered in a cellular environment, out of cell interference is also included in this noise term.

For the downlink, after joint transmission of the weighted signals bounded for different users from the base station, the baseband signal received by the mobile $i$ is given by

$$
\chi_i(t) = \sum_{k=1}^{K} \sqrt{P_k} b_k^H \sum_{l=1}^{L_k} F_{D_k}^D(t) \mathbf{a}^D(\theta_u) c_k(t - \tau_u) + n_i(t),
$$

(2.5)\\

where $w_k$ is a unit-norm transmit beamforming weight vector for user $k$, and $P_k$ is the power assigned to user $k$’s signal, which are the two parameters we want to design. $n_i(t)$ is a complex white Gaussian process. Due to the reciprocity of the uplink and downlink, other elements of the equation (2.5) are self-explanatory. As we noted above, although the uplink and downlink share many common features, the instantaneous fading coefficient and the steering vector are different for FDD system. To be specific, the downlink fading coefficients are given by

$$
F_{D_k}^D(t) = \frac{C}{d_{k1}^{1/2}} \sqrt{S_k} \alpha_{k1}^D(t),
$$

(2.6)
where all the large-scale parameters are identical to the uplink, but the Rayleigh fading is drawn from an independent instantiation. The downlink steering vector is given by

$$\mathbf{a}^D(\theta_k) = [1, e^{j2\pi d_L \frac{f_0}{c} \sin(\theta_k)}, \cdots, e^{-j2\pi d_L (M-1) \frac{f_0}{c} \sin(\theta_k)}]^T. \quad (2.7)$$

While the antenna elements should be closely spaced (e.g., half wavelength) for beamforming to get coherent signals across the antenna array, they should be widely separated (e.g., 10 wavelengths) to get diversity gain for transmit diversity schemes. Rather than being combined with a steering vector, the signals coming from different elements of antennas exploiting transmit diversity experience uncorrelated fading. Assume we have $M$ transmit antennas exploiting code transmit diversity, which transmit the data of $K$ users simultaneously. Then the transmitted signal from the $m$th antenna can be modeled as

$$s_m(t) = \frac{1}{\sqrt{M}} \sum_{k=1}^{K} \sqrt{P_k} b_k c_{km}(t), \quad (2.8)$$

where $c_{km}(t)$ is the spreading code for the $k$th user in the $m$th antenna with the spreading gain $N$, and the total transmitted energy of one user is normalized. Note that while array processing applies different weights for each antenna element on the signals of one user, (code) transmit diversity assigns different spreading codes for each antenna element to the signals of one user. The channel between the $m$th antenna and the $i$th receiver can be modeled as

$$h_m^i(t) = \sum_{l=1}^{L_i} F_m^i \delta(t - \tau_l^i), \quad (2.9)$$
where $F_{ml}^i$ is the complex path gain, and $\tau_l^i$ is the path delay. The received signal at the $i$th mobile is given by

$$r_i(t) = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} \sum_{k=1}^{K} \sqrt{P_k} b_k \sum_{l=1}^{L} F_{ml}^i c_{km}(t-\tau_l^i) + n_i(t). \quad (2.10)$$

### 2.2.2 FDD Framework

The essential element in antenna array beamforming design is the spatial covariance matrix, which, according to (2.5), is given by

$$R_{true,k}^D = \sum_{l} \sum_{l'} |F_{kl}^D|^2 |F_{kl'}^D|^2 a^D(\theta_{kl}) a^D(\theta_{kl'})^H. \quad (2.11)$$

Unfortunately, neither the instantaneous fading coefficients nor the steering vectors are known at the base stations, so estimation and approximation are necessary. Although the fading is uncorrelated between the uplink and downlink, their average strength is assumed to be insensitive to small changes in frequency [55], [82], i.e.,

$$E\{|F_{kl}^D|^2 |F_{kl'}^D|^2\} = E\{|F_{kl}^U|^2 |F_{kl'}^U|^2\}, \quad (2.12)$$

which can be estimated via time average from uplink data. To estimate the downlink steering vectors, several approaches exist. One idea (the matched array) is to design two separate closely located arrays which are scaled versions of each other in proportion to the ratio of the uplink and downlink wavelengths, thus making the uplink and downlink steering vectors the same [82]. The drawbacks of this approach are cost, imperfect array matching and near field uneven scattering. A clever log-periodic array configuration is proposed in [55], which overlaps the two subarrays of $M$ elements mentioned above into
one $M + 1$ array with $d_m / d_{m-1} = \lambda_u / \lambda_D$, where $d_m$ is the spacing between the $m$th and the $(m + 1)$th element. The drawbacks above are alleviated but still exist. Another approach (the duplex array) is to use a single array for both the uplink and downlink, and to transpose the array response from the uplink to the downlink via a linear transformation. However, some constraints are imposed to make the linear transformation tractable, e.g., a small frequency shift assumption in [82] and a circular array geometry in [3]. In our work, we exploit the approach to estimate the downlink array response from the uplink data through high-resolution DOA estimation methods or training sequences. We also ignore the estimation errors, an issue which deserves further study.

To sum up, we assume perfect knowledge of downlink direction of departure and calculate the array response for the downlink as (2.7). Then the approximated downlink spatial covariance matrix is given by

$$ R^D_{\text{approx},k} = \sum_{l} \sum_{l'} E[|F^U_{kl}||F^U_{kl'}|^2] \mathbf{a}^D(\theta_{kl}) \mathbf{a}^D(\theta_{kl'})^H. \quad (2.13) $$

### 2.2.3 Cellular System

We consider a cellular geometry as shown in Fig. 2.1. It consists of two tiers of surrounding cells around the cell of interest. Each cell is divided into three sectors of 120 degrees. Because we exploit CDMA, the sector of interest will suffer interference from adjacent sectors of the same cell, as well as out of cell, as indicated in Fig. 2.1. The out-of-cell interference will be assumed as white Gaussian and included in the noise term of the model, so only its power matters. The six-sector case is similar and thus is omitted here. We assume that all cells and sectors are identical. They are loaded with the same
number of users exhibiting the same behavior, and at the base stations the same operations are exploited. All the cells and sectors operate at the same time. This model should reflect the average performance of actual systems in the long run.

2.3 Power Control/Allocation Algorithms

In different application scenarios, optimal power control may have different meanings. A commonly used criterion is formulated as follows:

\[
\min \sum_{k=1}^{K} P_k \quad s.t. \quad \text{SINR}_k \geq \gamma_k, \quad 1 \leq k \leq K, \quad (2.14)
\]

i.e., minimize the total transmitted power with the constraints that each link obtains a SINR above a certain threshold. This is the optimization criterion we adopt for circuit-switched systems. For the packet-switched systems, we always transmit at the maximum
power, and we are concerned with the throughput of the network. We can simply allocate power equally among the active users, or we can allocate power in some optimal way. An optimal power assignment scheme proposed in [130] is formulated as follows:

\[
\text{max} \text{SINR}_{\min} \quad \text{s.t.} \quad \sum_{k=1}^{K} P_k \leq P_{\max},
\]

i.e., maximize the minimum link SINR with the total transmitted power constraint. This scheme tries to be fair to all users, which is not necessarily a good strategy for maximum throughput. It would be better to combine the study of physical layer transmission with the data link budget and network schedule, which is beyond the scope of this study.

Before we go into the details of the power control schemes, we first state a result underlying many such schemes.

### 2.3.1 Perron-Frobenius Theorem and its Applications

**Theorem:** Suppose \( T \) is a \( n \times n \) non-negative irreducible matrix. Then there exists an eigenvalue \( r \) such that:

a. \( r \) is real and positive;

b. \( r \) is associated with strictly positive left and right eigenvectors;

c. \( r = \max \{|\lambda_i|\} = \rho(T) \), where \( \lambda_i \), \( 1 \leq i \leq n \) are the eigenvalues of the matrix \( T \), and \( \rho(T) \) denotes its spectral radius;

d. \( r \) has algebraic multiplicity 1;

e. \( \min_i \sum_{j=1}^{n} t_{ij} \leq r \leq \max_i \sum_{j=1}^{n} t_{ij} \) with equality on either side implying equality throughout.

A similar result holds for column sums.
**Application 1:** A necessary and sufficient condition for a non-negative (non-trivial) solution \( \mathbf{x} \) to the equations \((s\mathbf{I} - \mathbf{T})\mathbf{x} = \mathbf{c}\) to exist for any nonnegative (non-trivial) vector \( \mathbf{c} \) is that \( s > r \). In this case there is only one strictly positive solution given by \((s\mathbf{I} - \mathbf{T})^{-1}\mathbf{c}\).

**Application 2:** If a non-negative (non-trivial) vector \( \mathbf{y} \) satisfies \( \mathbf{Ty} \leq s\mathbf{y} \), then

a) \( \mathbf{y} > 0; \)

b) \( s \geq r; \)

c) \( s = r \) if and only if \( \mathbf{Ty} = s\mathbf{y} \).

In the above, non-negative (\( \geq 0 \)) (strictly positive (\( > 0 \)), resp.) means that all the elements of a vector or matrix are nonnegative (strictly positive, resp.); and a trivial vector or matrix is one having all-zero elements.

### 2.3.2 General Form of Power Control Problem

For the sake of simplicity, we consider only the general form of power control here. In the next section, exact SINR formulas will be given and can be fit into this general setting without difficulty. The power control criterion of (2.14) is related to Application 1 of the theorem as follows. The general form of the power control problem is reformulated as

\[
\min \sum_{k=1}^{K} P_k \quad s.t. \quad (\mathbf{I} - \mathbf{DF})\mathbf{p} = \mathbf{u},
\]

(2.16)

where \( \mathbf{I} \) is a \( K \) by \( K \) identity matrix, \( \mathbf{D} \) is a diagonal matrix with entries \( \gamma_1, \cdots, \gamma_K \), \( \mathbf{F} \) is a non-negative irreducible matrix (interference term), \( \mathbf{p} = [P_1, P_2, \cdots, P_K]^T \) collects the powers assigned to all users, and \( \mathbf{u} \) is a positive vector (noise term). So we have a feasible
(positive) solution for power allocation vector if and only if the spectral radius of $\mathbf{DF}$ is less than one, otherwise we will claim an outage occurs. We call this a type-I outage and call the case in which we do get a positive solution but the total transmitted power exceeds a threshold, i.e., $\mathbf{p}^T \mathbf{1} > P_{\text{max}}$, a type-II outage. The solution to (2.16), if it exists, is given by $(\mathbf{I} - \mathbf{DF})^{-1} \mathbf{u}$ or alternatively by Jacobi iteration

$$\mathbf{p}^{(n+1)} = \mathbf{u} + \mathbf{DFp}^{(n)},$$  

(2.17)

which will converge for any initial value in this setting.

The power control criterion of (2.15) is related to Application 2 of the theorem as follows. It can easily be shown that this optimization scheme results in equal $\text{SINR} = \gamma$ for all links. The objective functions then become

$$\mathbf{p} = \gamma (\mathbf{Fp + h}) \text{ and } \mathbf{p}^T \mathbf{1} = P_{\text{max}}$$  

(2.18)

with $h_i = u_i / \gamma_i$. On writing $\mathbf{y} = [\mathbf{p}^T, 1]^T$, we can rewrite (2.18) as

$$\mathbf{T} \mathbf{y} = \gamma \mathbf{Qy}$$  

(2.19)

with

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{K \times K} & \mathbf{0}_{K \times 1} \\ \mathbf{1}^T & -P_{\text{max}} \end{bmatrix} \text{ and } \mathbf{Q} = \begin{bmatrix} \mathbf{F} & \mathbf{h} \\ \mathbf{0}_{1 \times K} & \mathbf{0} \end{bmatrix},$$  

(2.20)

where $\mathbf{1}$ is an all-1 vector. Alternatively, we can write it as

$$\mathbf{R} \mathbf{y} = \frac{1}{\gamma} \mathbf{y},$$  

(2.21)

with
It is easily shown that $\mathbf{R}$ is a non-negative irreducible matrix. So we always have a unique positive solution for $\mathbf{p}$ and the SINR margin is the reciprocal of the largest eigenvalue of $\mathbf{R}$.

### 2.4 Array Signal Processing

In this section, various array signal processing techniques are discussed in detail, among which are transmit diversity, sectorization, and beamforming techniques including beam steering, max SNR, and max SIR or SINR. We assume that the mobile receiver can learn the fading channel and perform RAKE combining. So the instantaneous SINR is obtained for each scheme, based on which the power control of Section 2.3 is then applied. A joint power control and beamforming algorithm is also discussed, and its optimality is verified.

#### 2.4.1 Transmit Diversity

We exploit code transmit diversity for downlink CDMA communications. The data streams of all users are transmitted simultaneously. For each user each data symbol is transmitted with equal power from every antenna using multiple mutually orthogonal spreading codes. A total of $KM$ Walsh codes are required in a straightforward design. To conserve codes, techniques such as space-time spreading can be used [54], however the $E_b/N_0$ performance achieved is no different than in the simple case above.
We can rewrite the discretized signal model of (2.8) as

\[ s_m = \frac{1}{\sqrt{M}} \sum_{k=1}^{K} \sqrt{P_k} b_k c_{km}. \]  

(2.23)

The channel between the \( m \)th antenna and the \( i \)th receiver of (2.9) is rewritten to emphasize the small-scale time-varying part

\[ h_m^i(t) = \sqrt{G_i} \sum_{l=1}^{L} \alpha_{ml}^i \delta(t - \tau_l^i), \]  

(2.24)

where \( G_i \) is the path gain from the transmit antennas to the \( i \)th user which combines the effect of path loss and shadowing, \( \alpha_{ml}^i \) is the instantaneous Rayleigh fading factor of the \( l \)th path from the \( m \)th antenna to the \( i \)th user, and \( \tau_l^i \) is the path delay. Although the antenna separation causes large variation of fading for different elements, the path delay is almost the same. If the path delay is within a few chips, we can ignore the intersymbol interference (ISI) and the received signal at the \( i \)th mobile is given by

\[ r_i = \frac{G_i}{\sqrt{M}} \sum_{m=1}^{M} \sum_{k=1}^{K} \sqrt{P_k} b_k C_{km} h_m^i + n_i, \]  

(2.25)

where \( C_{km} = [c_{km}^1, \ldots, c_{km}^K] \), which are delayed versions of \( c_{km} \); \( h_m^i = [\alpha_{m1}^i, \ldots, \alpha_{ml}^i]^T \); and \( n_i \) is the complex Gaussian noise. On denoting \( b = [b_1, \ldots, b_K]^T \), \( \tilde{l}^i_m = C_{km} h_m^i \), \( l'_i = \sum_{m=1}^{M} \tilde{l}^i_m \), \( L' = [l'_1, \ldots, l'_K] \), and \( P = \sqrt{G_i}/\sqrt{M} \) \( \text{diag}(\sqrt{P_1}, \ldots, \sqrt{P_K}) \), we have

\[ r_i = L' P b + n_i. \]  

(2.26)

A standard space-time RAKE receiver yields
\[
\text{Re}\{\mathbf{I}_i^H \mathbf{r}_i\} = \sqrt{\frac{G_i}{M}} \sqrt{P_i} \text{Re}\{(\mathbf{I}_i^')^H \mathbf{I}_i'\}b_i + \sum_{k \neq i} \frac{G_k}{M} \sqrt{P_k} \text{Re}\{(\mathbf{I}_k')^H \mathbf{I}_k'\}b_k + \text{Re}\{(\mathbf{I}_i')^H \mathbf{n}_i\}. \quad (2.27)
\]

Assuming that
\[
\left\{ \mathbf{e}^{(1)}_{k_1,m_1}, \mathbf{e}^{(2)}_{k_2,m_2} \right\} = \begin{cases} 
1 & l_1 = l_2, k_1 = k_2, m_1 = m_2 \\
0 & l_1 = l_2, (k_1,m_1) \neq (k_2,m_2), \\
\beta & l_1 \neq l_2
\end{cases}, \quad (2.28)
\]

where \( \beta \) is a random variable with
\[
E\{\beta\} = 0 \quad \text{and} \quad E\{\beta^2\} = \frac{1}{N}, \quad (2.29)
\]

and on denoting
\[
A^i = \sum_{m=1}^{M} \sum_{l=1}^{L} |\alpha_{ml}^i|^2 \quad (2.30)
\]

and
\[
C^i = \frac{1}{N} \sum_{m} \sum_{m'} \sum_{l} \sum_{l' \neq l} |\alpha_{ml}^i|^2 |\alpha_{m'l'}^i|^2, \quad (2.31)
\]

the SINR for user \( i \) is given by
\[
\text{SINR}_i = \frac{\frac{G_i}{M} P_i (A^i)^2}{\sum_{k \neq i} \frac{G_k}{M} P_k C^i + \sigma_i^2 A^i}, \quad (2.32)
\]

where \( \sigma_i^2 \) is the noise power at user \( i \)'s receiver.

The power control formula (2.16) is exemplified here with
\[
\mathbf{D} = \text{diag}(\gamma_1, \cdots, \gamma_K), \quad (2.33)
\]
\[ F_{ij} = \begin{cases} 0 & i = j \\ \frac{C_i}{NA} & i \neq j \end{cases} \]  
(2.34)

and

\[ u_i = \frac{\gamma_i \sigma_i^2 M}{G_i A_i} . \]  
(2.35)

### 2.4.2 Sectorization

The co-channel interference in a cellular system may be decreased by replacing omnidirectional antennas with directional antennas, each radiating within a specified sector. Sectorization usually increases users’ SINR or equivalently increases the system capacity, at the expense of increased number of antennas and decrease in trunking efficiency. The sectorizing antenna radiation pattern adopted here is formulated as follows [124]:

\[ G_s(\theta) = \begin{cases} 1 - \frac{(1-b)(\pi/S)^2}{(\pi/S)^2} \theta^2 & |\theta| \leq \sqrt{\frac{1-a}{1-b} \pi} \\ a & \text{elsewhere} \end{cases} \]  
(2.36)

![Fig. 2.2 Sector antenna radiation pattern](image-url)
where $G_s(\theta)$ is the gain of the antenna in a direction at angle $\theta$ relative to the maximum gain direction, $a$ denotes front-to-back ratio, $b$ denotes the attenuation at sector crossover, and $S$ is the number of sectors per cell. The antenna gain pattern for three and six sectors are given in Fig. 2.2 with $10 \log a = -15$ dB and $10 \log b = -3$ dB.

In our study, all techniques are employed in three-sector cells except the transmit diversity scheme, which is also studied in the six-sector cell case.

### 2.4.3 Beamforming Techniques

Unlike the transmit diversity situation, where antennas are separated far apart to get diversity gain, for beamforming, antennas are closely spaced so that signals coming at or going from the array elements are correlated. As the signals are coherent while the underlying noise is uncorrelated, antenna gain is obtained through judicious design of weighting vectors to combine (or pre-apply to) the antenna array signals. Before we discuss the various beamforming options, let us first assume generally a set of unit-norm transmit weighting vectors of $\{w_i\}_{i=1}^K$ are adopted for the $K$ users’ signals at the base stations. Again ignore ISI and separate the large-scale fading from the small-scale fading, the discretized signal model of (2.5) can be rewritten as

$$r_i = \sqrt{G_i} \sum_{k=1}^{K} \sqrt{P_k b_k} w_k^H \sum_{j=1}^{L} \alpha_{ij}^D a_{ij}^D c_j^l + n_i.$$ (2.37)

Denote $C_i = [c_i^1, \cdots, c_i^L]$, $h_i^D = [\alpha_{i1}^D, \cdots, \alpha_{iL}^D]^T$, and $l_i = C_i h_i^D$, then a standard space-time RAKE receiver of user $i$ yields (we ignore the superscript $D$ for simplicity)
\[ z_i = \mathbf{l}_i^H \mathbf{r}_i = \sqrt{P_i b_i} \mathbf{w}_i^H \sqrt{G_i} \left( \sum_{l=1}^{L} \sum_{l'} \alpha_{il}^* \alpha_{il} (\mathbf{c}_l^i)^H \mathbf{c}_l^i \mathbf{a}_d \right) + \sum_{k \neq i} \sqrt{P_k b_k} \mathbf{w}_k^H \sqrt{G_i} \left( \sum_{l=1}^{L} \sum_{l'} \alpha_{il}^* \alpha_{il} (\mathbf{c}_l^i)^H \mathbf{c}_l^i \mathbf{a}_d \right) + \sum_{l=1}^{L} \alpha_{il}^* (\mathbf{c}_l^i)^H \mathbf{n}_i. \] (2.38)

With the assumption of

\[ \langle \mathbf{c}_{k_1}^i, \mathbf{c}_{k_2}^i \rangle = \begin{cases} 1 & l_1 = l_2, k_1 = k_2, \\ 0 & l_1 = l_2, k_1 \neq k_2, \\ \beta & l_1 \neq l_2 \end{cases}, \] (2.39)

where \( \beta \) is a random variable defined in (2.29), and denoting

\[ R_i = G_i \sum_{j \neq i} |\alpha_{ij}|^2 |\alpha_{ij'}|^2 \mathbf{a}_d \mathbf{a}_d^H \] (2.40)

and

\[ Q_i = \frac{1}{N} G_i \sum_{j \neq i} |\alpha_{ij}|^2 |\alpha_{ij'}|^2 \mathbf{a}_d \mathbf{a}_d^H, \] (2.41)

the instantaneous SINR for user \( i \) is given by

\[ \text{SINR}_i = \frac{P_i \mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{k \neq i} P_k \mathbf{w}_k^H \mathbf{Q}_k \mathbf{w}_k + \sigma_i^2 \sum_{l=1}^{L} |\alpha_{il}|^2}. \] (2.42)

The power control formula (2.16) is exemplified here with

\[ D = \text{diag} \left( \frac{\gamma_1}{\mathbf{w}_1^H \mathbf{R}_1 \mathbf{w}_1}, \ldots, \frac{\gamma_k}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right), \] (2.43)

\[ F_{ij} = \begin{cases} 0 & i = j, \\ \mathbf{w}_j^H \mathbf{Q}_i \mathbf{w}_j & i \neq j, \end{cases} \] (2.44)

and
\[ u_i = \frac{\gamma_i \sigma_i^2 \sum_{l=1}^{L} |\alpha_{il}|^2}{w_i^H R_i w_i}. \]  

(2.45)

Since the downlink fading coefficients are not known at the base station, the approximation of Section 2.2.2 is adopted as follows:

\[ \mathbf{R}_i = G_i \sum_{l} \sum_{l'} E\{ |\alpha_{il}|^2 |\alpha_{il'}|^2 \} a_{il} a_{il'}^H, \]  

(2.46)

and

\[ \mathbf{Q}_i = \frac{1}{N} G_i \sum_{l} \sum_{l',l''} E\{ |\alpha_{il}|^2 |\alpha_{il'}|^2 \} a_{il} a_{il'}^H. \]  

(2.47)

Based on these matrices, various beamforming schemes are illustrated below. Equations (2.42) - (2.45) can be adjusted accordingly.

**2.4.3.1 Beam Steering**

This is a simple beamforming technique where the transmit antenna array forms a beam in the direction of line-of-sight of the desired user. This corresponds to the following antenna weights:

\[ w_i = \frac{a^D(\theta_{i,\text{los}})}{\|a^D(\theta_{i,\text{los}})\|}, \]  

(2.48)

where \( \theta_{i,\text{los}} \) denotes the azimuth angle of the line-of-sight of the \( i \)th user with the antenna array.
2.4.3.2 Maximum SNR

This scheme maximizes the SNR at the $i$th user. According to (2.42), it is equivalent to

$$\arg\max_{w_i} w_i^H \overline{R}_j w_i.$$  \hfill (2.49)

It is well known that the solution to (2.49) is given by the principal eigenvector of the matrix $\overline{R}_j$.

2.4.3.3 Maximum SIR/SINR

The Max SIR scheme transmits as much energy as possible to the desired user while minimizing its interference to other users. It is formulated as

$$\arg\max_{w_i} \frac{w_i^H \overline{R}_j w_i}{w_i^H \overline{T}w_i}, \quad \text{with} \quad \overline{T}_j = \sum_{k \neq i} \overline{Q}_k.$$  \hfill (2.50)

Such $w_i$ is given by the generalized principal eigenvector of $[\overline{R}_j, \overline{T}_j]$. Compared to Max SNR, this criterion may lead to inadequate power being transmitted to the desired user, or equivalently, may lead to increased transmitted power that results in a type-II outage. Intuitively, there is no benefit putting too much emphasis on interference minimization at the cost of reduced energy to the desired user, since the noise term cannot be eliminated.

In the packet-switched system, our goal is to maximize the network throughputs with the maximum transmit power, so the power allocation is known in advance. In this case, Max SINR can be exploited as follows:

$$\arg\max_{w_i} \frac{w_i^H \overline{R}_j w_i}{w_i^H \overline{T}w_i}, \quad \text{with} \quad \overline{T}_j = \sum_{k \neq i} \overline{Q}_k + \frac{K \sigma^2}{P_j} \mathbf{I},$$  \hfill (2.51)

which can be seen as a tradeoff between the Max SNR and Max SIR schemes.
2.4.4 Joint Power Control and Maximum SINR Beamforming

The beamforming approaches given in the last subsection are not the optimum downlink beamforming. While uplink beamforming is a decoupled problem (a chosen weight vector impacts only the desired receiver), in transmit beamforming each transmit weighting affects all the receivers. So downlink beamforming should be done jointly for all users.

The joint power control and beamforming problem was first considered and solved in part in [87], [88], where the uplink joint algorithm is proposed and proved to converge to the optimal solution, and a feasible solution is obtained for the downlink through virtual uplink construction. A complete solution to the joint optimal power control and downlink beamforming is given in [121] through normalization with the noise term (see (2.42)):

$$\tilde{R}_i = \frac{R_i}{\sum \alpha_i^2 \sigma_i^2} \quad \text{and} \quad \tilde{Q}_i = \frac{Q_i}{\sum \alpha_i^2 \sigma_i^2}. \quad (2.52)$$

The optimization problem is given by

$$\min_{\|w_i\|=1} \sum_{i=1}^{K} P_i \quad \text{s.t.} \quad \text{SINR}_i \geq \gamma_i \quad \text{and} \quad \|w_i\|=1, \quad (2.53)$$

with the SINR formula given by

$$\frac{P_i w_i^H \tilde{R}_i w_i}{\sum_{k \neq i} P_k w_k^H \tilde{Q}_i w_k + 1}. \quad (2.54)$$

The idea is to construct a virtual uplink problem with SINR

$$\frac{P_i w_i^H \tilde{R}_i w_i}{\sum_{k \neq i} P_k w_k^H \tilde{Q}_k w_i + \|w_i\|^2}. \quad (2.55)$$
The following iterations converge to the optimal beamforming vector and power allocation from any initial values for the virtual uplink problem (The superscript \( n \) means the \( n \)th iteration).

1. **Beamforming:** For \( 1 \leq i \leq K \),

\[
\mathbf{w}^n_i = \arg \max_{\mathbf{w}_i} \frac{\mathbf{w}_i^H \tilde{\mathbf{R}}_i \mathbf{w}_i}{\mathbf{w}_i^H \tilde{\mathbf{T}}_i^n \mathbf{w}_i},
\]  

(2.56)

where

\[
\tilde{T}^n_i = \sum_{k \neq i} (p^n_U)_k \tilde{Q}_k + \mathbf{I},
\]

(2.57)

with \( p^n_U = [(P^u_1)_n, \ldots, (P^u_K)_n]^T \) collecting the power at the \( n \)th iteration. This is the decentralized Max SINR scheme whose solution is the principal generalized eigenvector of \( [\tilde{\mathbf{R}}_i, \tilde{T}^n_i] \);

2. **Power control:**

\[
p^{n+1}_U = \tilde{\mathbf{D}}^n \tilde{\mathbf{F}}^n_U p^n_U + \tilde{\mathbf{u}}^n_U,
\]

(2.58)

where we define

\[
\tilde{\mathbf{D}}^n = \text{diag} \left( \frac{\gamma_1}{(w_1^n)^H \tilde{\mathbf{R}}_1 w_1^n}, \ldots, \frac{\gamma_K}{(w_K^n)^H \tilde{\mathbf{R}}_K w_K^n} \right)
\]

(2.59)

\[
(\tilde{\mathbf{F}}^n_U)_{ij} = \begin{cases} 
0 & i = j \\
(w_i^n)^H \tilde{Q} j w_j^n & i \neq j 
\end{cases}
\]

(2.60)
and

\[(\mathbf{u}^n_U)_i = \frac{\gamma_i}{\left(\mathbf{w}^n_i\right)^H \mathbf{R}_i \mathbf{w}^n_i} = \mathbf{D}^n \mathbf{1}_w, \quad (2.61)\]

where

\[\mathbf{1}_w = [\|\mathbf{w}_1\|, \ldots, \|\mathbf{w}_K\|]^T = \mathbf{1}. \quad (2.62)\]

This is the decentralized power control solution (see (2.17)) when the beamforming vector is fixed.

When the above algorithm converges, the optimal virtual uplink power vector is given by

\[\mathbf{p}_U = (\mathbf{I} - \mathbf{D}^n \mathbf{F}_U)^{-1} \mathbf{u}_U, \quad (2.63)\]

where \(\mathbf{D}, \mathbf{F}_U, \) and \(\mathbf{u}_U\) are converged values of (2.59), (2.60) and (2.61), respectively.

In line with (2.60) and (2.61), we define

\[(\mathbf{F}^n)^{ij} = \begin{cases} 0 & i = j \quad \text{(i.e., } \mathbf{F}_U^n = (\mathbf{F}^n)^T) \\ (\mathbf{w}_j^H \mathbf{Q}_i \mathbf{w}_j) & i \neq j \end{cases} \quad (2.64)\]

and

\[(\mathbf{u}^n)^{i} = \frac{\gamma_i}{\left(\mathbf{w}^n_i\right)^H \mathbf{R}_i \mathbf{w}^n_i} = \mathbf{D}^n \mathbf{1} = (\mathbf{u}^n_U)_i. \quad (2.65)\]

Then we claim that the optimum downlink power vector is given by

\[\mathbf{p} = (\mathbf{I} - \mathbf{D}^n \mathbf{F})^{-1} \mathbf{u}, \quad (2.66)\]
where $\mathbf{\tilde{D}}$, $\mathbf{\tilde{F}}$ and $\mathbf{\tilde{u}}$ are converged values of (2.59), (2.64) and (2.65), respectively. This is because

\[
\begin{equation}
1^T \mathbf{p} = 1^T (\mathbf{I} - \mathbf{\tilde{D}}\mathbf{\tilde{F}})^{-1} \mathbf{\tilde{D}} \mathbf{1} = 1^T \mathbf{\tilde{D}} (\mathbf{I} - \mathbf{\tilde{D}}\mathbf{\tilde{F}})^{-1} \mathbf{1} = (\mathbf{p}_{\text{ur}})^T \mathbf{1},
\end{equation}
\]

so the optimality of $\mathbf{p}$ is guaranteed by the optimality of the virtual uplink solution.

\subsection*{2.5 Numerical Results}

In this section, we examine the performance of the various downlink transmission techniques discussed above through computer simulation. For circuit-switched systems, power control is carried out and outage is declared whether there is no feasible power allocation to satisfy all the link requirements (type-I) or the total transmitted power exceeds the maximum value (type-II). We will evaluate and compare the performance of these downlink transmission techniques through the metric of total outage. In all cases, at least $\frac{10}{\text{outage}}$ simulations are run, and for outage $> 1\%$, 5000 simulations are run. For the packet-switched system, we allow each base station to transmit at the maximum power and equally divide the power between the active users. We examine the cumulative distribution function (CDF) of the SINR of a typical mobile user for performance comparison since the SINR is directly related to the achievable rate of the user. We also examine the effect of the optimal power assignment scheme of [130]. 5000 simulations are run for the packet-switched case.
In our setting, the maximum transmitted power to background noise ratio (out-of-cell interference not included) is set to be 30 dB, and we assume all the users see the same background noise level. The link SINR threshold is 5 dB for circuit-switched systems. The path loss parameter $\eta = 4$, and the standard deviation of the lognormal shadowing is 8 dB. The small-scale fading coefficients are generated through the typical urban (TUx) model used in W-CDMA 3G studies. The users are distributed uniformly within the sector of interest, with the antenna gain pattern given in Fig. 2.2. We assume each user has three multipaths, i.e., $L_1 = \cdots = L_k = L = 3$, the angles of which are Gaussian distributed around the direction of line-of-sight, with standard deviation of 10 degrees. The CDMA spreading gain is $N = 64$. The number of antennas $M$ in our study is 2, 4, or 8 per sector. We assume each cell has three 120-degree sectors unless otherwise noted. When studying the transmit diversity scheme in the six sector case, we assume the total number of users and antennas per cell is kept constant. So in the six-sector scenario, the number of users and antennas per sector is reduced to one half of those in the three-sector scenario.

### 2.5.1 Circuit-Switched System

Figures 2.3 to 2.5 present the performance of six transmission techniques combined with power control for CDMA downlink, namely, transmit diversity, transmit diversity with sectorization, Max SNR beamforming, Max SIR beamforming, beam steering, and joint power control and (Max SINR) beamforming. We assume no feedback from the mobile, while the loss due to this lack of feedback is also examined in the following. For the sake of comparison, the number of users that can be supported in one cell with 5% outage is given in Table 2.1.
Fig. 2.3 Performance comparison of various transmission techniques with $M = 2$ antennas per sector (6 antennas per cell) — circuit-switched system

Fig. 2.4 Performance comparison of various transmission techniques with $M = 4$ antennas per sector (12 antennas per cell) — circuit-switched system
Fig. 2.5 Performance comparison of various transmission techniques with $M = 8$ antennas per sector (24 antennas per cell) — circuit-switched system

### Table 2.1 Number of users supported in a cell with 5% outage

<table>
<thead>
<tr>
<th>Transmit Diversity</th>
<th>Six Sector Beam Steering</th>
<th>Max SNR</th>
<th>Max SNR with feedback</th>
<th>Joint with feedback</th>
<th>Joint with feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>36</td>
<td>39</td>
<td>42</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>36</td>
<td>57</td>
<td>51</td>
<td>66</td>
<td>72</td>
<td>66</td>
</tr>
<tr>
<td>39</td>
<td>69</td>
<td>12</td>
<td>81</td>
<td>117</td>
<td>84</td>
</tr>
</tbody>
</table>

From the above data, several conclusions can be made for CDMA downlink circuit-switched systems.
• We note that Max SNR beamforming approaches the optimal performance (that of joint power control and Max SINR beamforming) in the outage range of interest, while having much lower complexity.

• Max SIR has totally unacceptable performance and thus is omitted in Table 1. As we said before, putting too much emphasis on minimizing the interference to other users will hurt the desired energy; so more power has to be assigned to achieve the SINR threshold, resulting in type-II outage. Another problem of Max SIR is due to insufficient degrees of freedom the antenna array can offer compared to the number of users.

• Beam steering has good performance only when the number of antenna elements are small ($M = 2$); the gap between beam steering and Max SNR beamforming enlarges as $M$ increases.

• For transmit diversity, sectorization significantly improves the performance (6 to 30 users more as $M$ goes from 2 to 8 at 5% outage); but the Max SNR beamforming technique still outperforms the six-sector transmit diversity scheme (6 to 12 users more as $M$ goes from 2 to 8 at 5% outage).

Figures 2.6 to 2.10 show the performance of four transmission techniques as the number of antennas per sector varies from 2, 4 to 8. Max SIR beamforming is not of interest due to its unacceptable performance. The performance of joint power control and beamforming is similar to that of Max SNR and omitted here. From these figures, the following conclusions can be drawn.
Fig. 2.6 Performance of transmit diversity with 2, 4 and 8 antennas per sector

Fig. 2.7 Performance of transmit diversity with sectorization with 2, 4 and 8 antennas per sector
Fig. 2.8 Performance of max SNR beamforming without feedback with 2, 4 and 8 antennas per sector

Fig. 2.9 Performance of max SNR beamforming with feedback with 2, 4 and 8 antennas per sector
Fig. 2.10 Performance of beam steering with 2, 4 and 8 antennas per sector

- For transmit diversity, the gain from exploiting more antennas diminishes as the number of antennas increases.

- For Max SNR beamforming the gain through exploiting more antenna elements does not diminish but is restricted due to imperfect channel knowledge. On the other hand, if we assume ideal feedback, the gain through exploiting more antenna elements increases with the number of antennas.

- For beam steering we observe an interesting phenomenon: the performance improves from $M = 2$ to $M = 4$, but deteriorates as $M$ further increases. One possible explanation is that the beam steering scheme forms a beam toward the physical position of the mobile. Due to the angle spread model we use, it points in the wrong directions. As more antennas are used, more precise calibration of the line-of-sight
actually means greater angle estimation errors. This effect will counteract the benefit of antenna gains with more antennas.

Finally, Figs 2.11 to 2.13 compare the Max SNR beamforming with and without feedback with $M = 2, 4$ and 8. Again those of Max SIR scheme are not of interest and those of joint power control and beamforming scheme are similar and are omitted. We note that the gap between that of no feedback and that with feedback increases as the number of antennas increases, but for small numbers of antennas ($M = 2, 4$), the loss due to approximation of channel parameters is insignificant. This means that for small number of antennas, Max SNR beamforming is the best choice even without feedback information.

![Fig. 2.11 Performance of max SNR beamforming with 2 antennas per sector: with and without feedback channel information](image)
Fig. 2.12 Performance of max SNR beamforming with 4 antennas per sector: with and without feedback channel information

Fig. 2.13 Performance of max SNR beamforming with 8 antennas per sector: with and without feedback channel information
2.5.2 Packet-Switched System

As a counterpart to the circuit-switched case, Figs 2.14 to 2.16 present the performance of six transmission techniques for packet-switched system. Because we do not have power control in this case, we study the optimal power assignment combined with Max SNR instead. We perform equal power assignment unless otherwise noted. For the sake of comparison, the mean (50% CDF) and peak (90% CDF) SINR values of a typical user are given in Table 2.2 and Table 2.3, respectively. Note that for the $M = 8$ and $K = 4$ case, the simultaneously transmitted users are doubled. One should consider this when translating SINR to achievable rates and network throughput. “(f)” in the tables means with feedback channel parameter information.

![Performance comparison of various transmission techniques with $M = 4$ antennas and 2 Active users — packet-switched system](image)

**Fig. 2.14** Performance comparison of various transmission techniques with $M = 4$ antennas and 2 Active users — packet-switched system
Fig. 2.15 Performance comparison of various transmission techniques with $M = 8$ antennas and 2 active users — packet-switched system

Fig. 2.16 Performance comparison of various transmission techniques with $M = 8$ antennas and 4 active users — packet-switched system
From the above data, several conclusions can be drawn for CDMA downlink packet-switched systems.

- Optimal power allocation scheme has no benefit in packet-switched systems. We see from Figs 2.14 to 2.16 that, the Max SNR with optimal power allocation, compared with Max SNR with equal power allocation, favors low-rate (low SINR) users but

<table>
<thead>
<tr>
<th>Transmit Diversity</th>
<th>Six Sector</th>
<th>Beam Steering</th>
<th>Max SNR</th>
<th>Max SNR(f)</th>
<th>Max SINR</th>
<th>Opt PA &amp; Max SNR</th>
<th>Opt PA &amp; Max SNR(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>11.8</td>
<td>10.8</td>
<td>11</td>
<td>11.1</td>
<td>10.6</td>
<td>10.6</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>11.4</td>
<td>12.4</td>
<td>12.7</td>
<td>13.2</td>
<td>13.7</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.3</td>
</tr>
<tr>
<td>4.2</td>
<td>7.2</td>
<td>7.6</td>
<td>8.5</td>
<td>8.8</td>
<td>6.4</td>
<td>7.2</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transmit Diversity</th>
<th>Six Sector</th>
<th>Beam Steering</th>
<th>Max SNR</th>
<th>Max SNR(f)</th>
<th>Max SINR</th>
<th>Opt PA &amp; Max SNR</th>
<th>Opt PA &amp; Max SNR(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>17.5</td>
<td>15.5</td>
<td>15.7</td>
<td>15.7</td>
<td>21.6</td>
<td>21.6</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.8</td>
<td>17.2</td>
<td>19.3</td>
<td>19.5</td>
<td>19.5</td>
<td>24.3</td>
<td>24.6</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17.1</td>
</tr>
<tr>
<td>6.8</td>
<td>10</td>
<td>12.7</td>
<td>13.3</td>
<td>13.4</td>
<td>18.8</td>
<td>19.3</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.6</td>
</tr>
</tbody>
</table>
harm high-rate (high SINR) users. As we discussed in Section 2.3, optimal power allocation is like a socialist scheme that seeks absolute fairness. It cannot achieve the highest throughput and, without being jointly considered with the link and network schedules, cannot guarantee fairness either. A similar phenomenon can be observed for the Max SINR scheme and is omitted here.

- Contrary to the circuit case, Max SINR beamforming has the best performance in terms of peak rate; it is also good at mean rate with small numbers of users, while comparable with others when there are more users.
- Max SNR beamforming is almost the best in terms of mean rate; it is also good in terms of peak rate performance.
- Beam steering is almost as good as max SNR in terms of peak rate performance, while a little worse (1 dB) in terms of mean rate performance.
- For transmit diversity, sectorization significantly improves the performance (4-7 dB).
- The Max SNR beamforming technique outperforms the six-sector transmit diversity scheme for $M = 8$ (1 dB in mean and 2-3 dB in peak); for $M = 4$, six-sector transmit diversity is better.

Figures 2.17 to 2.19 compare Max SNR beamforming with and without feedback, while Figs. 2.20 to 2.22 show the Max SINR case. We find that feedback does not help much for beamforming in packet-switched systems.
Fig. 2.17 Performance of max SNR beamforming with 4 antennas and 2 active users: with and without feedback channel information

Fig. 2.18 Performance of max SNR beamforming with 8 antennas and 2 active users: with and without feedback channel information
Fig. 2.19 Performance of max SNR beamforming with 8 antennas and 4 active users: with and without feedback channel information.

Fig. 2.20 Performance of max SINR beamforming with 4 antennas and 2 active users: with and without feedback channel information.
Fig. 2.21 Performance of max SINR beamforming with 8 antennas and 2 active users: with and without feedback channel information.

Fig. 2.22 Performance of max SINR beamforming with 8 antennas and 4 active users: with and without feedback channel information.
2.6 Summary

In this chapter, we have seen that traffic type impacts the algorithm choice in downlink beamforming. For circuit-switched downlink CDMA systems, the Max SNR beamforming scheme is the best choice (accommodating 12 to 42 more users than transmit diversity). For packet-switched systems, Max SINR Beamforming has the best performance in terms of peak rate (10-14 dB more than transmit diversity); Max SNR beamforming is almost the best in terms of mean rate (3-4 dB more than transmit diversity), but beam steering and transmit diversity with sectorization are also good choices. We also see that sectorization greatly improves the system performance, both for the circuit and for the packet case. For transmit diversity, the gain from exploiting more antennas diminishes, while this is not the case for beamforming, especially with feedback channel information. The following issues deserve further study in this context.

- In this chapter, we have assumed the perfect knowledge of DOA when calculating the downlink spatial covariance matrix. The issue of parameter estimation errors in covariance matrix calculation deserves further study.
- As we have noted, the optimal power allocation has no benefit on the packet-switched systems. Combining link and network schedules with transmission techniques in packet-switched systems is of interest.
- In this chapter, we have mainly studied array-processing techniques to combat the small-scale fading. Their counterparts for widely separated antennas (macrodiversity) are also of interest.
Chapter 3

Receive Arrays:
Iterative Space-Time Multiuser Detection

3.1 Introduction

In this chapter, space-time multiuser detection for wireless communications with receive arrays is studied. To overcome the computational burden that rises very quickly with increasing numbers of users and receive antennas in asynchronous multipath CDMA channels, efficient implementations of space-time multiuser detection algorithms, including batch iterative methods and sample-by-sample adaptive methods, are considered here.

The presence of both multiple-access interference (MAI) and intersymbol interference (ISI) constitutes a major impediment to reliable high-data-rate CDMA communications in multipath channels. These phenomena present challenges as well as opportunities for receiver designers: through multiuser detection (MUD) [119] and space-time (ST) processing [78], the inherent code, spatial, temporal and spectral diversities of multipath multi-antenna CDMA channels can be exploited to achieve substantial gain.
Advanced signal processing typically improves system performance at the cost of computational complexity. It is well known that the optimal maximum likelihood (ML) multiuser detector has prohibitive computational requirements for most current applications. A variety of linear and nonlinear multiuser detectors have been proposed to ease this computational burden while maintaining satisfactory performance. However, in asynchronous multipath CDMA channels with receive antenna arrays and large data frame lengths, direct implementation of these suboptimal techniques still proves to be very complex. Techniques for efficient space-time multiuser detection fall largely into two categories. One includes batch iterative methods, which assume knowledge of all signals and channels and is suitable, for example, for base station processing in cellular systems. The other includes sample-by-sample adaptive methods, which require knowledge only of the signal and (possibly) channel of a desired user. Sample-by-sample adaptive methods are suitable both for mobile end processing, which entails decentralized data detection, and for base station processing due to the time varying nature of wireless communications. We may think of the sample-by-sample adaptive methods discussed here as being most suitable for application at the base station, where it is more practical to install an antenna array. However, most of the described decentralized or blind techniques are readily applied to the mobile user end when multiple antennas can be applied at mobile terminals.

This chapter is organized as follows. In Section 3.2 a space-time multiuser signal model is presented. Batch iterative methods are discussed in Section 3.3 while sample-by-sample adaptive methods are dealt with in Section 3.4. Numerical examples are given.
in Section 3.3 and 3.4 to show the performance of the various iterative ST MUD techniques developed in this chapter. Section 3.5 summarizes this chapter.

### 3.2 Space-Time Signal Model

Consider a direct-sequence CDMA communication system with $K$ users employing normalized spreading waveforms $s_1, \ldots, s_K$ given by

$$s_k(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} c_k(j) \psi(t - jT_c), \quad 0 \leq t \leq T, \quad 1 \leq k \leq K,$$  \hspace{1cm} (3.1)

where $N$ is the processing gain, \{ $c_k(j); \quad 0 \leq j \leq N-1$ \} is a signature sequence of $\pm1$’s assigned to the $k$th user, and $\psi(\cdot)$ is a normalized chip waveform of duration $T_c = T/N$ with $T$ the symbol interval. User $k$ (for $1 \leq k \leq K$) transmits a frame of $M$ independent equiprobable BPSK symbols $b_k(i) \in \{+1, -1\}, \quad 0 \leq i \leq M - 1$; and the symbol sequences from different users are assumed to be mutually independent. The transmitted baseband signal due to the $k$th user is thus given by

$$x_k(t) = A_k \sum_{i=0}^{M-1} b_k(i) s_k(t - iT), \quad 1 \leq k \leq K,$$  \hspace{1cm} (3.2)

where $A_k$ is the amplitude associated with user $k$’s transmission. The transmitted signal of each user passes through a multipath channel before it is received by a uniform linear antenna array (ULA) of $P$ elements with inter-element spacing $d$. Then the single-input multiple-output (SIMO) vector impulse response between the $k$th user and the receive array can be modeled as
\[ h_k(t) = \sum_{l=1}^{L} a_{kl} g_{kl} \delta(t - \tau_{kl}), \quad (3.3) \]

where \( L \) is the maximum number of resolvable paths between each user and the receive array (for simplicity we assume \( L \) is the same for each user), \( g_{kl} \) and \( \tau_{kl} \) are respectively the complex gain and delay of the \( l \)th path of the \( k \)th user, and

\[
a_{kl} = \begin{bmatrix}
a_{kl,1} \\
a_{kl,2} \\
\vdots \\
a_{kl,p}
\end{bmatrix} = \begin{bmatrix}
1 \\
e^{j2\pi d \sin(\theta_{kl})/\lambda} \\
\vdots \\
e^{j2\pi d (P-1) \sin(\theta_{kl})/\lambda}
\end{bmatrix} \quad (3.4)
\]

is the ULA response corresponding to the signal of the \( l \)th path of the \( k \)th user with direction of arrival (DOA) \( \theta_{kl} \) and carrier wavelength \( \lambda \). \( \delta(t) \) denotes the Dirac delta function. The received signal at the antenna array is the superposition of the channel-distorted signals from the \( K \) users together with additive Gaussian noise, which is assumed to be spatially and temporally white. This leads to the vector received signal model

\[ r(t) = \sum_{k=1}^{K} x_k(t) \otimes h_k(t) + \sigma n(t), \quad (3.5) \]

where \( \otimes \) denotes convolution, and \( \sigma^2 \) is the spectral height of the ambient Gaussian noise at each antenna element.

A sufficient statistic for demodulating the multiuser symbols from the space-time signal (3.5) is given by [127]

\[ y = [y_1(0), \ldots, y_E(0), y_1(1), \ldots, y_1(M-1), \ldots, y_E(M-1)]^T, \quad (3.6) \]
where the elements \{y_k(i)\} are defined as follows:

\[
y_k(i) = \sum_{l=1}^{L} g_{kl}^* a_{kl}^H \int_{\tau_{zl}}^{\tau_{zl+1}} r(t) s_k(t-iT-\tau_{zl}) dt, \quad 1 \leq k \leq K, \quad 0 \leq i \leq M-1.
\] (3.7)

To produce this sufficient statistic, the received signal vector \(r(t)\) is first match-filtered for each path of each user to form the vector observables \(\{z_{zl}(i)\}\), after which beams are formed on each path of each user via the dot products with the array responses \(\{a_{zl}\}\), and then all the paths of each user are combined with a RAKE receiver. This process produces one observable for each symbol of each user. Since the system is in general asynchronous and the users are not orthogonal, we need to collect the statistic for all users over the entire data frame. The observable \(y_k(i)\) corresponds to the output of a conventional space-time matched filter, matched to the \(i\)th symbol of user \(k\). Therefore, a general space-time multiuser receiver is (as shown in Fig. 3.1) a space-time matched filter bank, followed by a decision algorithm. In the following, we will present various ST MUD receivers based on this space-time matched filter output. However, a new ST MUD receiver structure will also be introduced, in which chip-level observables are exploited.

The sufficient statistic (3.6) can be written as (see [119])

\[
y = HAb + \sigma v,
\] (3.8)

where \(H\) is a \(KM \times KM\) matrix capturing the cross-correlations between different symbols and different users, \(A\) is the \(KM \times KM\) diagonal matrix whose \(k+iK\) diagonal elements are equal to \(A_k\), \(b = [b_1(0), \ldots, b_k(0), b_1(1), \ldots, b_1(M-1), \ldots, b_k(M-1)]^T\), and \(\sigma\) is the standard deviation.
v \sim \mathcal{N}(0, \mathbf{H}) \text{ (i.e., v is Gaussian with zero mean and covariance matrix } \mathbf{H}). \text{ An optimal ML space-time multiuser detector will maximize the following log-likelihood function }

\Omega(\mathbf{b}) = 2 \text{Re}\{\mathbf{b}^\text{T} \mathbf{y}\} - \mathbf{b}^\text{T} \mathbf{A} \mathbf{H} \mathbf{A} \mathbf{b} \ . \hspace{1cm} (3.9)

\text{Fig. 3.1 A conventional space-time multiuser receiver structure}

The multiuser signal and channel parameters (signature waveforms, multipath delay and amplitude, array response) come into play through the $KM \times KM$ block Toeplitz system matrix $\mathbf{H}$, which can be written as

$$
\mathbf{H} \triangleq 
\begin{bmatrix}
H^{[0]} & H^{[1]} & \cdots & H^{[\Delta]}

H^{[-1]} & H^{[0]} & H^{[1]} & \cdots & H^{[\Delta]}

H^{[-\Delta]} & \cdots & H^{[0]} & \cdots & H^{[\Delta]}

H^{[-\Delta]} & \cdots & H^{[-1]} & \cdots & H^{[0]}

H^{[-\Delta]} & \cdots & H^{[-1]} & \cdots & H^{[0]}
\end{bmatrix} \ . \hspace{1cm} (3.10)
$$
where $\Delta$ denotes the multipath delay spread, and $H^{[i]} = (H^{[i]})^{H}$. The $n, m$th element of $H$ is the cross-correlation between the composite received signatures (after beamforming and RAKE combining) of the $n$th and $m$th elements of $b$. The reader is referred to [127] for further details of $H$. Dynamic programming can be applied to compute the ML estimates of $b$. Due to the binary nature of $b$, the complexity of this computation is on the order of $O(2^{(\Delta+1)K} / K)$ per user per symbol.

### 3.3 Batch Iterative Methods

There has been considerable research in space-time processing (e.g., [67], [78]), most of which considers single-user-based methods. Combined multiuser detection and array processing has been addressed recently (e.g. [72], [127]). In this section, we consider iterative implementation of linear and nonlinear space-time multiuser detectors (ST MUD) in multipath CDMA channels with receive antenna arrays. In particular, we develop several such algorithms, and compare them on the basis of performance and complexity. Ultimately, we conclude that an algorithm based on the expectation-maximization (EM) algorithm offers an attractive tradeoff in this context.

#### 3.3.1 Iterative Linear ST MUD

In this part, we consider the application of iterative processing to the implementation of various linear space-time multiuser detectors in algebraic form. After the introduction to the general form of linear ST MUD, we go on to discuss two general approaches to
iteratively solving large systems of linear equations. Subsequent subsections will treat nonlinear iterative methods.

Linear multiuser detectors in the framework of (3.8) are of the form

\[
\hat{b} = \text{sgn}(\text{Re}\{Wy\}),
\]

where \( W \) is a \( KM \times KM \) matrix. For the linear decorrelating (zero-forcing) detector, this matrix is given by

\[
W_d = H^{-1},
\]

while for the linear MMSE detector, we have

\[
W_m = (H + \sigma^2 A^{-2})^{-1}.
\]

Direct inversion of the matrices in (3.12) and (3.13) (after exploiting the block Toeplitz structure) is of complexity \( O(K^2M\Delta) \) per user per symbol [48].

The linear multiuser detection estimates of (3.11) can be seen as the solution of a linear equation

\[
Cx = y,
\]

with \( C = H \) for the decorrelating detector and \( C = H + \sigma^2 A^{-2} \) for the MMSE detector. Jacobi and Gauss-Seidel iteration are two common low-complexity iterative schemes for solving linear equations such as (3.14) [48]. If we decompose the matrix \( C \) as \( C = C_L + D + C_U \) where \( C_L \) denotes the lower triangular part, \( D \) denotes the diagonal part, and \( C_U \) denotes the upper triangular part, then Jacobi iteration can be written as

\[
x_m = -D^{-1}(C_L + C_U)x_{m-1} + D^{-1}y,
\]
and Gauss-Seidel iteration is represented as

\[ x_m = - (D + C_L)^{-1} C_U x_{m-1} + (D + C_L)^{-1} y. \] \hspace{1cm} (3.16)

From (3.15), Jacobi iteration can be seen to be a form of linear parallel interference cancellation [89], the convergence of which is not guaranteed in general. One of the sufficient conditions for the convergence of Jacobi iteration is that \( D - (C_L + C_U) \) be positive definite [48]. In contrast, Gauss-Seidel iteration, which (3.16) reveals to be a form of linear serial interference cancellation, converges to the solution of the linear equation from any initial value, under the mild conditions that \( C \) is symmetric and positive definite [48], which is always true for the MMSE detector.

Another approach to solving the linear equation (3.14) involves gradient methods, among which are steepest descent and conjugate gradient iteration [48]. The reader is referred to Section 3.4 for sample-by-sample adaptive space-time processing methods, which apply gradient methods in a different setting. Note that solving (3.14) is equivalent to minimizing the cost function

\[ \Phi(x) = \frac{1}{2} x^H C x - x^H y. \] \hspace{1cm} (3.17)

The idea of gradient methods is the successive minimization of this cost function along a set of directions \( \{p_m\} \) via

\[ x_m = x_{m-1} + \alpha_m p_m, \] \hspace{1cm} (3.18)

with

\[ \alpha_m = p_m^H q_{m-1} / p_m^H C p_m, \] \hspace{1cm} (3.19)
and

\[ \mathbf{q}_m = -\nabla \Phi(x) \bigg|_{x=x_m} = \mathbf{y} - \mathbf{C}x_m. \] (3.20)

Different choices of the set \( \{ \mathbf{p}_m \} \) in (3.18) - (3.20) give different algorithms. If we choose the search directions \( \mathbf{p}_m \) to be the negative gradient of the cost function \( \mathbf{q}_{m-1} \) directly, this algorithm is the steepest descent method, global convergence of which is guaranteed [48]. The convergence rate may be prohibitively slow, however, due to the linear dependence of the search directions, resulting in redundant minimization. If we choose the search direction to be \( \mathbf{C} \)-conjugate as follows

\[ \mathbf{p}_m = \arg\min_{\mathbf{p} \in \Lambda_{m-1}} \| \mathbf{p} - \mathbf{q}_{m-1} \|, \] (3.21)

where \( \Lambda_m = \text{span} \{ \mathbf{Cp}_1, ..., \mathbf{Cp}_m \} \), then we have the conjugate gradient method, whose convergence is guaranteed and performs well when \( \mathbf{C} \) is near the identity either in the sense of a low rank perturbation or in the sense of norm [48]. The computational complexity of Gauss-Seidel and conjugate gradient iteration are similar, which is on the order of \( O(K \Delta \bar{m}) \) per user per symbol, where \( \bar{m} \) is the number of iterations.

### 3.3.2 Iterative Nonlinear ST MUD

Nonlinear multiuser detectors are often based on bootstrapping techniques, which are iterative in nature. In this part, we will consider the iterative implementation of decision-feedback multiuser detection in the space-time domain. We also discuss briefly the implementation of multistage interference cancelling ST MUD, which serves as a
reference point for introducing a new EM-based iterative ST MUD, to be discussed in the next subsection.

### 3.3.2.1 Cholesky Iterative Decorrelating Decision-Feedback ST MUD

Decorrelating decision feedback multiuser detection (DDF MUD) [34], [117] exploits the Cholesky decomposition \( \mathbf{H} = \mathbf{F}^H \mathbf{F} \), where \( \mathbf{F} \) is a lower triangular matrix, to determine feedforward and feedback matrices for detection via the algorithm

\[
\hat{\mathbf{b}} = \text{sgn}(\mathbf{F}^{-H} \mathbf{y} - (\mathbf{F} - \text{diag}(\mathbf{F})) \mathbf{A} \mathbf{b}),
\]

(3.22)

which should be understood to detect the bits sequentially from that of the first user to the last user, with \( \hat{\mathbf{b}} \) a vector containing the detected bits for all users over the whole data frame.

Suppose the user of interest is user \( k \) (each bit of each user can be treated as a “new” user for asynchronous systems), the purpose of the feedforward matrix \( \mathbf{F}^{-H} \) is to whiten the noise and decorrelate against the “future users” \( \{s_{k+1}, \ldots, s_{KM}\} \); while the purpose of the feedback matrix \( (\mathbf{F} - \text{diag}(\mathbf{F})) \) is to cancel out the interference from “previous users” \( \{s_1, \ldots, s_{k-1}\} \). Note that the performance of DDF MUD is not uniform. While the first user is demodulated by its decorrelating detector, the last detected user will essentially achieve its single-user lower bound providing the previous decisions are correct. There is another form of Cholesky decomposition, in which the feedforward matrix \( \mathbf{F} \) is upper triangular. If we were to use this form instead in (3.22), then the multiuser detection would be in the reverse order, as would be the performances. The idea of *Cholesky iterative DDF ST*
MUD is to employ these two forms of Cholesky decomposition alternatively as follows. For lower triangular Cholesky decomposition $F_1$, first feedforward filtering is applied as

$$\bar{y}_1 = F_1^{-H} y, \quad (3.23)$$

where it is readily shown that $\bar{y}_{1,i} = F_{1,ii} A_i b_i + \sum_{j=1}^{i-1} F_{1,ij} A_j b_j + \bar{n}_{1,i}, \quad i = 1, \ldots, KM$, with $\bar{n}_{1,i}, i = 1, \ldots, KM$, the independent and identically distributed (i.i.d.) Gaussian noise components with zero mean and variance $\sigma^2$. We can see that the influence of the “future users” is eliminated and the noise component is whitened. Then we employ the feedback filtering to take out the interference from “previous users” as

$$u_1 = \bar{y}_1 - (F_1 - \text{diag}(F_1)) \hat{A} \hat{b}, \quad (3.24)$$

where it is easily seen that $u_{1,i} = \bar{y}_{1,i} - \sum_{j=1}^{i-1} F_{1,ij} A_j \hat{b}_j \approx F_{1,ii} A_i b_i + \bar{n}_{1,i}, \quad i = 1, \ldots, KM$. Similarly, for upper triangular Cholesky decomposition $F_2$, we have

$$\bar{y}_2 = F_2^{-H} y, \quad (3.25)$$

where $\bar{y}_{2,i} = F_{2,ii} A_i b_i + \sum_{j=i+1}^{KM} F_{2,ij} A_j b_j + \bar{n}_{2,i}, \quad i = KM, \ldots, 1$, and

$$u_2 = \bar{y}_2 - (F_2 - \text{diag}(F_2)) \hat{A} \hat{b}, \quad (3.26)$$

where $u_{2,i} = \bar{y}_{2,i} - \sum_{j=i+1}^{KM} F_{2,ij} A_j \hat{b}_j \approx F_{2,ii} A_i b_i + \bar{n}_{2,i}, \quad i = KM, \ldots, 1$.

After the above operations are (alternately) executed, the following log-likelihood ratio is calculated,
\[ L_i = 2 \text{Re}(\mathbf{F}_{1/2,i}^* A_i u_{1/2,i}) / \sigma^2, \quad (3.27) \]

where \( \mathbf{F}_{1/2} \) and \( u_{1/2} \) are used to give a shorthand representation for both alternatives.

Then the log-likelihood ratio is compared with the last stored value, which is replaced by the new value if the new one is more reliable, i.e.,

\[
L_i^{\text{stored}} = \begin{cases} 
L_i^{\text{stored}} & \text{if } |L_i^{\text{stored}}| > |L_i^{\text{new}}| \\
L_i^{\text{new}} & \text{otherwise}
\end{cases} \quad (3.28)
\]

Finally, we make soft decisions \( \hat{b}_i = \tanh(L_i / 2) \) at an intermediate iteration, which has been shown to offer better performance than making hard intermediate decisions, and make hard decisions \( \hat{b}_i = \text{sgn}(L_i) \) at the last iteration. Three or four iterations are usually enough for the system to achieve an improved steady state without significant oscillation. The structure of Cholesky iterative decorrelating decision-feedback ST MUD is illustrated in Fig. 3.2.

**Fig. 3.2** Cholesky iterative decorrelating decision-feedback ST MUD
The Cholesky factorization of the block Toeplitz matrix $\mathbf{H}$ (see (3.10)) can be done recursively. For $\Delta = 1$

\[
\mathbf{F} = \begin{bmatrix}
\mathcal{F}(0) & 0 & 0 & 0 \\
\mathcal{F}(0) & \mathcal{F}(0) & \cdots & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \cdots & \mathcal{F}_M(1) & \mathcal{F}_M(0)
\end{bmatrix},
\]

where the element matrices are obtained recursively as follows:

\[
\mathcal{V}_M = \mathcal{H}^{[0]},
\]

and, for $i = M, M - 1, \ldots, 1$, we do the Cholesky decomposition for the reduced-rank matrix $\mathcal{V}_i$ to get $\mathcal{F}_i(0)$

\[
\mathcal{V}_i = \mathcal{F}_i^H(0)\mathcal{F}_i(0),
\]

while $\mathcal{F}_i(1)$ is obtained as

\[
\mathcal{F}_i(1) = (\mathcal{F}_i^H(0))^{-1}\mathcal{H}^{[-1]}.
\]

Finally we have

\[
\mathcal{V}_{i-1} = \mathcal{H}^{[0]} - \mathcal{H}^{[1]}\mathcal{V}_i^{-1}\mathcal{H}^{[-1]}
\]

for use in the next iteration. The extension of this algorithm to $\Delta > 1$ is straightforward and is omitted here.
3.3.2.2 Multistage Interference Cancelling ST MUD

Multistage interference cancellation (IC) [115] is similar to Jacobi iteration except that hard decisions are made at the end of each stage in place of the linear terms that are fed back in (3.15). Thus we have

\[
\hat{b}_m = \text{sgn} \left( y - (C_L + C_U)\hat{b}_{m-1} \right) = \text{sgn} \left( y - (H - D)\hat{b}_{m-1} \right).
\] (3.34)

The underlying rationale of this method is that the estimator-subtractor structure exploits the discrete-alphabet property of the transmitted data streams. This nonlinear hard-decision operation typically results in more accurate estimates, especially in high SNR situations. Although the optimal decisions are a fixed point of the nonlinear transformation (3.34), there are problems with the multistage IC such as a possible lack of convergence and oscillatory behavior. In the following section we consider some improvement over space-time multistage IC MUD.

Except for the Cholesky factorization, the computational complexity for Cholesky iterative DDF ST MUD is the same as multistage IC ST MUD, which is essentially the same as that of linear interference cancellation, i.e., \( O(K\Delta\tilde{n}) \) per user per symbol.

3.3.3 EM-based Iterative ST MUD with a New Structure

In this part, EM and space-alternating generalized EM (SAGE) algorithms are applied to the space-time multipath asynchronous CDMA systems to avoid the convergence and stability problem of the multistage IC MUD. To address the problems caused by long spreading codes, namely, the large computational burden to obtain the cross-correlation matrix \( H \), a new space-time multiuser receiver structure is also introduced. The SAGE
algorithm is then applied nontrivially to group spatial-domain multiuser detection based on different directions of arrival for different paths of different users. The SAGE iterative ST MUD with this new structure retains its excellent performance while with the conventional structure but with greater adaptability; it is easily adjusted for use in time-varying environments and for non-CDMA space-time processing.

3.3.3.1 EM and SAGE Algorithm with Application to ST MUD

The EM algorithm [30] provides an iterative solution of maximum likelihood estimation problems such as

$$
\hat{\theta}(Y) = \arg \max_{\theta \in \Lambda} \log f(Y; \theta),
$$

(3.35)

where $\theta \in \Lambda$ are the parameters to be estimated, and $f(\cdot)$ is the parameterized probability density function of the observable $Y$. The idea of the EM algorithm is to consider a judiciously chosen set of “missing data” $Z$ to form the complete data $X = \{Y, Z\}$ as an aid to the parameter estimation, and then to iteratively maximize the following new objective function

$$
Q(\theta; \bar{\theta}) = E\{\log f(Y, Z; \theta) | Y = y; \bar{\theta}\},
$$

(3.36)

where $\theta$ are the parameters in the likelihood function, which are to be estimated, while $\bar{\theta}$ represent a priori estimates of the parameters from the previous iteration. Together with the observations these previous estimates are used to calculate the expected value of the log-likelihood function with respect to the complete data $X = \{Y, Z\}$. To be specific, given an initial estimate $\theta^0$, the EM algorithm alternates between the following two steps:
(a). E-step, where the complete-data sufficient statistic $Q(\theta; \theta^i)$ is computed;

(b). M-step, where the estimates are refined by $\theta^{i+1} = \arg \max_{\theta \in \Lambda} Q(\theta; \theta^i)$.

It is well known that EM estimates monotonically increase the likelihood, and converge stably to an ML solution under certain conditions [30].

An issue in using the EM algorithm is the tradeoff between ease of implementation and convergence rate. One would like to add more “missing data” to make the complete data space more informative so that the implementation of the EM algorithm is simpler than the original setting (3.35). However, the convergence rate of the algorithm is inversely proportional to the Fisher information contained in the complete data space [38]. Thus, the convergence rate of the EM algorithm is notoriously slow, especially for multidimensional parameter estimation, due to the simultaneous updating nature of the M-step of the EM algorithm. The SAGE algorithm has been proposed in [38] to improve the convergence rate for multidimensional parameter estimation. The idea is to divide the parameters into several groups (subspaces), with only one group being updated at each iteration. Thus, we can associate multiple less-informative “missing data” sets to improve the convergence rate while maintaining overall tractability of optimization problems. For each iteration, a subset of parameters $\theta_{S_i}$ and the corresponding missing data $Z_{S_i}^\perp$ are chosen, which is called the definition step. Then similar to the EM algorithm, in the E-step we calculate

$$Q^S(\theta_{S_i}; \theta^i) = E\{\log f(Y, Z^S; \theta_{S_i}, \theta^i_{S_i} | Y = y; \theta^i)\},$$

(3.37)

where $\theta_{S_i}$ denotes the complement of $\theta_{S_i}$ in the whole parameter set; in the M-step, the chosen parameters are updated while the others remain unchanged as
\[
\begin{aligned}
\theta_{S_i}^{i+1} &= \arg \max_{\theta \in \Lambda_{S_i}} Q^S_i(\theta_{S_i}; \theta^i) \\
\theta_{S_i}^{i+1} &= \theta_{S_i}^i
\end{aligned}
\tag{3.38}
\]

where \( \Lambda_{S_i} \) denotes the restriction of the entire parameter space to those dimensions indexed by \( S_i \). Like the traditional EM estimates, the SAGE estimates also monotonically increase the likelihood and converge stably to an ML solution \[38\].

The EM algorithm can be applied to space-time multipath asynchronous CDMA multiuser detection as follows. For ease of illustration, we reindex the vectors and matrices in the system model (3.8) as

\[
y = [y_1, \ldots, y_K, y_{K+1}, \ldots, y_{K(M-1)+1}, \ldots, y_{KM}]^T,
\]

\[
b = [b_1, \ldots, b_K, b_{K+1}, \ldots, b_{K(M-1)+1}, \ldots, b_{KM}]^T,
\]

and

\[
A = \text{diag}[a_1, \ldots, a_K, a_{K+1}, \ldots, a_{K(M-1)+1}, \ldots, a_{KM}]^T.
\]

Suppose we would like to detect a bit \( b_k, k \in \{1, 2, \cdots, KM\} \), while the interfering users’ bits \( b_j = \{b_j\}_{j \neq k} \) are treated as the missing data. The complete-data sufficient statistic is given by (\( H_{km} \) is the element of matrix \( H \) at the \( k \)th row and \( m \)th column)

\[
Q(b_k; \tilde{b}_k) = \frac{a_k^2}{2\sigma^2} \left( -b_k^2 + \frac{2b_k}{a_k} (y_k - \sum_{m \neq k} H_{km} a_m \tilde{b}_m) \right),
\tag{3.39}
\]

with

\[
\tilde{b}_m = E\{b_m \mid Y = y; b_k = \tilde{b}_k\} = \tanh\left( \frac{a_m}{\sigma^2} (y_m - H_{mk} a_k b_k) \right),
\tag{3.40}
\]
which forms the E-step of the EM algorithm. The M-step is given by

$$\begin{equation}
    b_{k}^{i+1} = \arg \max_{b_k \in \Lambda} Q(b_k; b_i^i) = \begin{cases} 
    \text{sgn}(y_k - \sum_{m \neq k} H_{km}a_m \tilde{b}_m) & \Lambda = \{\pm 1\} \\
    \frac{1}{a_k}(y_k - \sum_{m \neq k} H_{km}a_m \tilde{b}_m) & \Lambda = \Re 
    \end{cases}
\end{equation}$$

(3.41)

where \(\Lambda = \Re\) (the set of real numbers) means a soft decision is needed, e.g., in an intermediate stage. Note that in the E-step (40), interference from users \(j \neq k\) is not taken into account, since these are treated as “missing data”. This shortcoming is overcome by the application of the SAGE algorithm, where the bit vector of all users \(b = \{b_j\}_{j=1}^{KM}\) is treated as the parameter to be estimated and no missing data is needed. The algorithm is described as follows: for \(i = 0, 1, \cdots\),

**Definition step:** \(S_i = 1 + (i \mod KM)\)

**M-step:**

\[
\begin{cases}
    b_{k}^{i+1} = \text{sgn}(y_k - \sum_{m \neq k} H_{km}a_m b^i_m) & k \in S_i \\
    b_{m}^{i+1} = b^i_m & m \notin S_i
\end{cases}
\]

Note that there is no E-step since there is no missing data, and interference from all other users is recreated from previous estimates and subtracted. The resulting receiver is similar to the multistage interference cancelling multiuser receiver (see (3.34)), except that the bit estimates are made sequentially rather than in parallel. However, with this simple concept of sequential interference cancellation, the resulting multiuser receiver is convergent, guaranteed by the SAGE algorithm. The multistage interference cancelling multiuser receiver discussed in 3.3.2.2, on the other hand, does not always converge. The computational complexity of this SAGE iterative ST MUD is also \(O(K \Delta \bar{m})\) per user per symbol.
3.3.3.2 SAGE Iterative ST MUD with a New Structure

So far when we discuss computational complexity, we have ignored the computational burden for calculating the matrix $H$. In a CDMA system exploiting short spreading codes as we have assumed, the system matrix will exhibit the block-Toeplitz structure as shown in (10). The computation of $H^{[j]}$, $1 \leq j \leq \Delta$, though involved (please see [127] for details), would become insignificant when $M$ is large. This is not the case, however, when a long spreading code is employed so that spreading sequences vary from symbol to symbol (e.g. in IS-95). In this situation, the sub-matrices $H^{[j]}$, $1 \leq j \leq \Delta$, have to be calculated for each symbol, which results in an additional complexity of $O(K^3L\Delta)$ per user per symbol.

![Diagram of a new space-time multiuser receiver structure](image)

**Fig. 3.3** A new space-time multiuser receiver structure
To circumvent this problem, a new space-time processing structure is introduced, where we separate the spatial and temporal processing, and apply the SAGE algorithm to spatial-domain multiuser detection based on different directions of arrival for different paths of different users, as shown in Fig. 3.3.

From (3.5), we can write

\[
\mathbf{r}(t) = \sum_{i=0}^{M-1} \sum_{k=1}^{K} A_k b_k(i) \sum_{l=1}^{L} \mathbf{a}_{kl} g_{kl} s_k(t - iT - \tau_{kl}) + \mathbf{n}(t). \tag{3.43}
\]

After chip-matched filtering and chip-rate sampling, the discrete time signal is given as

\[
\mathbf{z}(n) = \int_{nT_c + \tau_{\min}}^{(n+1)T_c + \tau_{\min}} \mathbf{r}(t) \psi(t - nT_c - \tau_{\min}) dt, \quad 0 \leq n \leq MN - 1 + \left\lceil \frac{\tau_{\max} - \tau_{\min}}{T_c} \right\rceil, \tag{3.44}
\]

where \( \tau_{\min} \) and \( \tau_{\max} \) denote the minimum and maximum value, respectively, of the delay profile for all users. For simplicity, we assume the delays are integer numbers of chip intervals. (The results can straightforwardly be extended to the fractional delay case by oversampling.) On denoting

\[
s_{kl}(n) = \int_{nT_c + \tau_{\min}}^{(n+1)T_c + \tau_{\min}} \mathcal{S}_{kl}(t - iT - \tau_{kl}) \psi(t - nT_c - \tau_{\min}) dt, \tag{3.45}
\]

we have

\[
\mathbf{z}(n) = \sum_{k=1}^{K} \sum_{l=1}^{L} A_k b_k \left( \left\lceil \frac{n + \tau_{kl} - \tau_{\min}}{T_c} \right\rceil \right) \mathbf{a}_{kl} g_{kl} s_{kl}(n) + \mathbf{\sigma e}(n) \tag{3.46}
\]

\[
= \sum_{j} \mathbf{\tilde{a}}_j \mathbf{\tilde{d}}_j(n) + \mathbf{\sigma e}(n), \quad 1 \leq j \leq K \times L,
\]
where for the last equality we reformulate the system model so that chips received from different paths of each user are treated as different users in a synchronous system, according to the following translation

$$\tilde{A}_j \leftrightarrow A_k g_{k \ell}, \quad \tilde{a}_j \leftrightarrow a_{k \ell}, \quad \text{and} \quad d_j \leftrightarrow b_k s_{k \ell}, \quad \text{(3.47)}$$

with $j = (k-1)L + l$, $1 \leq k \leq K$, $1 \leq l \leq L$. Here $e(n)$ denotes i.i.d. white background noise with zero mean and unit variance. The parameters to be estimated are

$$\theta = \mathbf{d} = [d_1(n), \ldots, d_{KL}(n)]^T \in \{-1\}^{KL}$$

which correspond to the chips from all paths of all users. The index sets cycle through $1, \ldots, KL$, with $L$ chips (of a user) being updated at a time. The algorithm is implemented without any missing data so there is no need for the E-step. Each iteration thus comprises the following steps:

Definition step: $S_i = [1, \ldots, L] + (i \mod K) \times L$;

M-step:

$$\tilde{d}_{j}^{i+1} = g_{A}\left(\tilde{a}_j^H(z - \sum_{m \neq j} \tilde{a}_m \tilde{a}_m^H d_m^i)\right) \quad j \in S_i \quad \text{(3.48)}$$

$$d_{m}^{i+1} = d_m^i, \quad m \notin S_i$$

Since the spatial processing is one of the components of our proposed space-time receiver, whose outputs are provided to the next stage for temporal processing, we choose to use the linear function $g_A(x) = x$ in the above M-step to produce soft outputs $\tilde{d}_j(n)$, $j \in S_i$, which is better for overall performance than the hard-decision function $g_A(x) = \text{sgn}(x)$. 
After the spatial processing described above, the chips from the different paths of one user are combined through a RAKE combiner to get a chip estimate for that user $(k = (i \mod K))$, 

$$
\hat{d}_k(n) = \sum_{l=1}^{L} A_k^* g_{kl} \tilde{d}_{kl}(n + \tau_{kl} / T_c), \quad 0 \leq n \leq MN - 1 + \left\lfloor \frac{\tau_{\text{max}} - \tau_{\text{min}}}{T_c} \right\rfloor,
$$

(3.49)

where the index conversion is made through the translation of (3.47). Finally, the spreading code is employed to get the bit estimate

$$
\hat{b}_k(i) = \text{sgn} \left( \text{Re} \left( \sum_{n=1}^{N} \hat{d}_k^n(iT + n)c_k(n) \right) \right), \quad 0 \leq i \leq M - 1.
$$

(3.50)

The obtained bit estimates are then respread and remodulated as (3.46) and (3.47) to get $d_j^{i+1}$, $j \in S_i$ for spatial domain interference cancellation (3.48) for the next iteration.

Our proposed space-time multiuser receiver structure in Fig. 3.3 has several advantages. As we mentioned earlier, the multiuser signals and multipath channel parameters come into play through the system matrix $H$. For a CDMA system employing long spreading codes, $H$ must be calculated for each symbol interval, which is quite cumbersome. More generally, for any time-varying communication system, $H$ has to be updated on the order of the coherence time. Furthermore, any part of the algorithm that is related to this system matrix, e.g., the Cholesky factorization for Cholesky iterative DDF ST MUD, should also be calculated for each symbol. This problem is circumvented by our structure, which effectively distributes the operations combined in the system matrix $H$ into the different stages of spatial IC, beamforming, temporal RAKE combining, and despreading. Further, since the front end processing is at the chip level, these “users” are
synchronous. Therefore, the algorithm can be implemented chip-by-chip or symbol-by-symbol in a pipelined version. This structure also has the benefit that it can be applied in non-CDMA (e.g. space-division multiple-access (SDMA)) cellular networks to exploit the spatial and temporal knowledge to suppress interference and improve system capacity. All we need in such cases is to omit the despreading stage.

The computational complexity of this new-structure-based SAGE iterative ST MUD is $O(KNL\bar{m})$ per user per symbol, which can easily be implemented with $O(KL\bar{m})$ or even less time complexity per user per symbol with modern VLSI techniques. The basic idea would be to build multiple parallel hardware processing units, which can be done due to the synchronous chip-level processing nature of the new structure. The multistage IC ST MUD can also be implemented with this structure in a straightforward way. The performance of these ST MUD receivers with this new structure is the same as that implemented with the conventional structure of Fig. 3.1, which will be discussed further in the next section.

### 3.3.4 Numerical Results

In this section, the performance of the above described space-time multiuser detectors is examined through computer simulations. We assume a $K = 8$-user CDMA system with spreading gain $N = 16$. Each user travels through $L = 3$ paths before it reaches a ULA with $P = 3$ elements and half-wavelength spacing. The maximum delay spread is set to be $\Delta = 1$. The complex gains and delays of the multipath and the directions of arrival are randomly generated and kept fixed for the whole data frame. This corresponds to a slow fading situation. The spreading codes of all users are randomly generated and kept fixed
for all the simulations. We assume $A_1 = \ldots = A_K$ for simplicity, but the received signal powers of different users are unequal due to the effects of multipath.

First we compare the performance of various space-time multiuser receivers and some single-user space-time receivers in Fig. 3.4. Five receivers are considered: the single-user matched filter, the single-user MMSE receiver, the multiuser MMSE receiver implemented in Gauss-Seidel or conjugate gradient iteration method (the performance is the same for both), the Cholesky iterative decorrelating decision-feedback multiuser receiver, and the multistage interference cancelling multiuser receiver. All these receivers are implemented on the conventional space-time multiuser receiver structure shown in Fig. 3.1. (The reader is referred to [127] for derivations of the single-user based receivers.) The performance is evaluated after the iterative algorithms converge. Due to the bad convergence behavior of the multistage IC MUD, we measure its performance after three stages. The single-user lower bound is also depicted for reference. We can see that the multiuser approach greatly outperforms the single-user based methods; nonlinear MUD offers further gain over the linear MUD; and the multistage IC seems to approach the optimal performance (not always, e.g., see Fig. 3.4(c)), when it has good convergence behavior. Note that due to the introduction of spatial (receive antenna) and spectral (RAKE combining) diversity, the SNR for the same BER is substantially lower than that required by normal receivers without these signal processing methods.

Figure 3.5 shows the performance of Cholesky iterative decorrelating decision-feedback ST MUD for two users, which is also typical for other users. Note that we use a different parameter set for this simulation, so there is no correspondence between Fig. 3.4 and Fig. 3.5. We find that the Cholesky iterative method offers uniform gain over its non-
iterative counterpart. This gain may be substantial for some users and negligible for others due to the individual characteristics of signals and channels.

In Section 3.3.3.2, a new space-time multiuser receiver structure was introduced to reduce complexity and enhance adaptability of the algorithms while keeping the same performance. In this new structure, different paths of each user are treated as separate users in a synchronous system, while in the original structure, different paths are combined prior to interference cancellation. One may be concerned that the new structure is more susceptible to saturation (“dimensional crowding”) than the traditional structure by a factor of $L$. We implemented the SAGE ST MUD with both structures for a fully loaded $K = N = 16$ system. The results shown in Fig. 3.6 are measured after three iterations. It is seen that the performance is almost identical for these two structures. We
CHAPTER 3. RECEIVE ARRAYS: ITERATIVE SPACE-TIME MULTIUSER DETECTION

User #5

Bit error rate vs. SNR (dB)

User #7

Bit error rate vs. SNR (dB)
Fig. 3.4 Performance comparison of BER versus SNR for five space-time multiuser receivers
believe that this should always be the case as long as the number of paths $L$ is no larger than the number of antenna elements $P$.

Finally, we show the advantage of our new EM-based (SAGE) iterative method over the multistage IC method with regard to the convergence of the algorithms. These receivers are implemented on the new space-time multiuser receiver structure shown in Fig. 3.3, and long random spreading codes are employed. We assume again a $K = 8$-user CDMA system with spreading gain $N = 16$. From Fig. 3.7 we find that, while the multistage interference cancelling ST MUD converges slowly and exhibits oscillatory behavior, the SAGE ST MUD converges quickly and outperforms the multistage IC method. The oscillation of the performance of the multistage IC corresponds to a performance degradation as no statistically best iteration number can be chosen.
Fig. 3.6 Performance comparison of SAGE ST MUD with traditional and new ST structure
Fig. 3.7 Performance comparison of convergence behavior of multistage interference cancelling ST MUD and EM-based iterative ST MUD
3.4 Sample-by-Sample Adaptive Methods

In contrast to batch iterative methods, the observation vectors of interest for sample-by-sample adaptive methods are not outputs of a space-time matched-filter bank but the chip-sampled signal itself. Thus, the cyclostationary character of interfering signals is preserved, which is essential for their removal by adaptive methods. To be specific, supposing the user of interest is the $k$th user, during the $i$th symbol interval the received signal at the $p$th antenna element is passed through a chip-matched filter and then sampled at the chip rate to obtain an $\mathbf{N}$-vector of signal samples

$$
\mathbf{r}_p^{(k)}(i) = [r_{p,0}^{(k)}(i), r_{p,1}^{(k)}(i), \ldots, r_{p,N-1}^{(k)}(i)]^T,
$$

where $\mathbf{N} = N + \lceil (\tau_{kL} - \tau_{k1}) / T_c \rceil$ (Without loss of generality, we assume $\tau_{k1} \leq \cdots \leq \tau_{kL}$ here.) is large enough to capture all the information of the desired user from all paths, and the samples are given by

$$
\int_{iT+(m+1)T_c+\tau_{k1}}^{iT+\tau_{k1}} r_p(t)\psi(t-iT-mT_c-\tau_{k1})dt.
$$

Then the $P\mathbf{N}$-vector $\mathbf{r}^{(k)}(i) = [\mathbf{r}_1^{(k)}(i)]^T, \ldots, [\mathbf{r}_p^{(k)}(i)]^T]^T$ becomes the sufficient statistic for the detection of $b_\delta(i)$ in the various space-time sample-by-sample adaptive receivers discussed in the sequel. Henceforth, we will omit the time index $i$ when no ambiguity is incurred.

3.4.1 Data-Aided ST MUD

Multiuser detectors can be divided into two categories: centralized and decentralized, depending on whether the knowledge of interfering users is required or not. For mobile-
end processing, decentralized or blind techniques are preferred since information relative to the other users is either difficult to obtain or forbidden. On the contrary, at the base station, centralized receivers can be exploited to improve the system’s performance, especially in situations with severe near-far problems and/or with great ISI. In this subsection, we will describe a decentralized adaptive MMSE space-time multiuser detector and a centralized adaptive decision-feedback space-time multiuser detector, both of which operate with the aid of training sequences.

### 3.4.1.1 Decentralized Adaptive MMSE ST MUD

Figure 3.8 depicts the structure of a decentralized adaptive space-time multiuser detector of interest in detecting user $k$’s $i$th symbol. Each antenna element is equipped with a chip-matched filter followed by a chip-interval-spaced adaptive finite-impulse-response (FIR) filter. The outputs of all FIR filters are summed and sampled at the symbol rate to form a soft decision output, which serves two purposes: to form an estimate for the desired bit through a decision device, and to form an error signal for adjustment of adaptive filter coefficients.

Collect the weights of the FIR filter banks at the $p$th antenna element into an $N$-vector $\mathbf{w}_p^{(k)} = [w_{p,0}^{(k)}, w_{p,1}^{(k)}, \cdots, w_{p,N-1}^{(k)}]^T$, and then collect such vectors from all antenna elements into a $P\tilde{N}$-vector $\mathbf{W}_k = [(\mathbf{w}_1^{(k)})^T, (\mathbf{w}_2^{(k)})^T, \cdots, (\mathbf{w}_p^{(k)})^T]^T$. $\mathbf{W}_k$ is thus applied to the signal vector $\mathbf{r}^{(k)}$ given in (3.51) to make a decision about $b_k$. A useful performance metric for the receiver of Fig. 3.8 is the output signal-to-noise ratio (SINR), which can be estimated as
\[ S\text{INR}_k = 10 \log_{10} \left( \frac{1 - MSE(W_k)}{MSE(W_k)} \right), \]  
\[ (3.53) \]

with mean square error defined as

\[ MSE(W_k) = E \left\{ (b_k - (W_k)^H r^{(k)})^2 \right\} \]
\[ = 1 - (W_k)^H h_k - (h_k)^H W_k + (W_k)^H R_k W_k, \]  
\[ (3.54) \]

where \( R_k = E\{r^{(k)} r^{(k)H}\} \) is the autocorrelation matrix of the signal vector \( r^{(k)} \), and \( h_k = E\{r^{(k)} b_k\} \) is the crosscorrelation vector between \( r^{(k)} \) and the desired bit \( b_k \). An optimum choice for \( W_k \) is that which minimizes the mean square error \( MSE(W_k) \). This choice, known as the MMSE detector, is given by the Wiener-Hopf solution

\[ W_k^{opt} = R_k^{-1} h_k. \]  
\[ (3.55) \]

For this theoretical optimum solution, the achieved minimum value of the mean square error is given by

\[ Fig. 3.8 \text{ Structure of an adaptive MMSE space-time multiuser detector} \]
\[ MMSE_k^\Delta = \text{MSE}(W_k^{\text{opt}}) = E \left\{ \left( b_k - (W_k^{\text{opt}})^H r^{(k)} \right)^2 \right\} \]
\[ = 1 - (W_k^{\text{opt}})^H R_k W_k^{\text{opt}} = 1 - (h_k)^H R_k^{-1} h_k. \] (3.56)

A number of algorithms are available to seek the solution (3.55) adaptively, from the simple least-mean-squares (LMS) algorithm to various fast yet complex recursive-least-squares (RLS) methods. The properties and behavior of these algorithms are well known and documented [53]. Here we adopt the LMS algorithm as a simple tool to obtain MMSE FIR filter banks. This choice is illustrated as follows. The soft decision output is given by
\[ y_k(i) = W_k^H(i) r^{(k)}(i), \] (3.57)
from which a bit estimate is formed as
\[ \hat{b}_k(i) = \text{sgn}\{\text{Re}(y_k(i))\}, \] (3.58)
where “Re” indicates the real part. An error signal is then formed as
\[ \varepsilon_k(i) = b_k(i) - y_k(i), \] (3.59)
and the filter coefficients are updated as
\[ W_k(i + 1) = W_k(i) + \mu \varepsilon_k^*(i) r^{(k)}(i), \] (3.60)
where \( \mu \) is the step size of the adaptive algorithm. Note that after the training period, the receiver is switched to decision-directed mode and the error signal is formed as
\[ \varepsilon_k(i) = \hat{b}_k(i) - y_k(i). \] (3.61)
It is well known [53] that the mean square error \( \text{MSE}_k(i) = |\varepsilon_k(i)|^2 \) converges to a steady state value if and only if the step size parameter satisfies the following two conditions:
0 < \mu < \frac{2}{\lambda_{\text{max}}}, \quad (3.62)

and

\sum_i \frac{\mu \lambda_i}{2 - \mu \lambda_i} < 1, \quad (3.63)

where \{\lambda_i\} are the eigenvalues of the autocorrelation matrix \( R_k \) and \( \lambda_{\text{max}} = \max_i \lambda_i \). If (3.62) is satisfied, (3.63) can be replaced by a weaker form

\[
\mu < \frac{2}{\sum_i \lambda_i} = \frac{2}{\text{tr}(R_k)} = \frac{2}{P_t},
\]

where \( P_t \) is the total input power. After fulfilling these two conditions, larger step size speeds up convergence but results in more excess MSE, while smaller step size causes the opposite effect. For adaptive interference suppression in CDMA systems it has been shown that a step size near \( 1/P_t \) gives the best convergence speed with a reasonable excess MSE [73]. Another commonly used method is to employ a variable step size: using a larger step size at the start for fast convergence and shrinking it as the system approaches the steady state for smaller excess MSE. The steady state MSE in the LMS algorithm is given by

\[
\text{MSE}_{\lambda}(\infty) = \frac{\text{MMSE}_k}{1 - \sum_i \frac{\mu \lambda_i}{(2 - \mu \lambda_i)}}. \quad (3.65)
\]
3.4.1.2 Centralized Adaptive Decision-Feedback ST MUD

Figure 3.9 gives the structure for a centralized adaptive decision-feedback ST MUD. In contrast with the structure of Fig. 3.8, in Fig. 3.9 previously detected bits of all active users are exploited through a symbol-spaced FIR feedback filter to help detect the bit of interest $b_k(i)$. The length of the feedback filter can be taken as the maximum delay spread (more than one symbol), while that of the chip-spaced feedforward filter can just span one symbol. As noted previously, centralized decision-feedback receivers are suitable for base station processing, and are expected to achieve great improvement over decentralized MMSE receivers in the case of large delay spread and severe near-far problem.

![Fig. 3.9 Structure of an adaptive centralized decision-feedback space-time multiuser detector](image)

Now the input signal is augmented to include the previous bits from all users; i.e., we consider $r_{n}^{(k)}(i) = [r_1^{(k)}(i)]^T, ..., [r_p^{(k)}(i)]^T, [e_1(i)]^T, ..., [e_K(i)]^T]^T$ with
where $D$ is the maximum delay spread in units of symbol interval. Again, collect the feedforward filter coefficients at the $p$th antenna element into the vector $\mathbf{w}_p = [w_{p,1}, w_{p,2}, \ldots, w_{p,N}]^T$, $1 \leq p \leq P$, and the feedback filter coefficients for the $n$th user with the vector $\mathbf{v}_n = [v_{n,1}, \ldots, v_{n,D}]^T$, $1 \leq n \leq K$. Collect all these feedforward and feedback filter vectors into a $(PN + KD)$-vector $\mathbf{W}_k = [\mathbf{w}_1^T, \mathbf{w}_2^T, \ldots, \mathbf{w}_P^T, \mathbf{v}_1^T, \ldots, \mathbf{v}_K^T]^T$. Detection then proceeds by applying this vector to the augmented input signal. The adaptation and analysis of the algorithm readily follows that in subsection 3.4.1.1.

### 3.4.1.3 Numerical Results

In this subsection, the performance of the above described data-aided adaptive space-time multiuser detectors is examined through computer simulations. We assume a $K = 16$-user CDMA system with spreading gain $N = 16$, which is heavily loaded with severe near-far problem. Each user travels through $L = 3$ paths before it reaches a ULA with $P = 3$ elements and half-wavelength spacing. The maximum delay spread is set to be $4T$. The complex gains and delays of the multipath and the directions of arrival are randomly generated and kept fixed for all the simulations. We assume $A_1 = \ldots = A_K$ for simplicity, but the received signal powers of different users are unequal due to the effects of multipath. The number of symbols per frame is $M = 250$. The step size of the LMS algorithm is fixed to be $\mu = 0.001$. 

$$e_n(i) = \left[ b_n(i-1), \ldots, b_n(i-D) \right]^T, \quad 1 \leq n \leq K,$$ (3.66)
Figure 3.10 shows the learning curve for the decentralized adaptive MMSE ST MUD. The user of interest is user 1. The theoretical MMSE is also plotted in the figure (as the dashed line) for comparison. Figure 3.11 compares the steady-state bit error rate (BER) of the decentralized adaptive MMSE ST MUD with that of the batch iterative MMSE ST MUD. The error is counted and averaged for consecutive 400 data frames after an initial 4 data frames (1000 iterations) of adaptation. These results show that this adaptive ST MUD structure approaches the optimum MMSE ST MUD, while using only knowledge of the timing and training sequence of the desired user. This simple adaptive structure effectively combines the function of beamforming, RAKE combining and multiuser detection.
Fig. 3.11 Bit error rate of the decentralized adaptive MMSE space-time multiuser detector in the steady state

Figures 3.12 and 3.13 compare the performance of the decentralized adaptive MMSE ST MUD and the centralized adaptive decision-feedback ST MUD. The user of interest is user 3 (a weak user with strong interference). Figure 3.12 compares their steady state output SINRs, which is given by (3.53) using $MSE_{k}(\infty)$ (see (3.65)). Figure 3.13 compares their steady state BERs. Both figures clearly demonstrate the significant improvement of the centralized adaptive decision-feedback ST MUD over the decentralized adaptive MMSE ST MUD in the situation of large delay spread and severe near-far problem.
Fig. 3.12 Comparison of steady state output SINR of the two adaptive receivers

Fig. 3.13 Comparison of steady state BER of the two adaptive receivers
3.4.2 Blind ST MUD

We see from the previous subsection that data aided adaptive space-time multiuser detectors achieve very good performance with simple algorithms, while no side information is needed. However, these receivers require training sequences from the desired user(s), which consume system resources. Moreover, whenever there is a dramatic change in the interference environment (e.g., a deep fade or the powering on of a strong interferer), decision directed adaptation becomes unreliable, and new training sequences have to be transmitted. Therefore, implementation of data aided adaptive algorithms in practical wireless CDMA networks, which favor complete asynchronous and uncoordinated transmissions that turn on and off autonomously, may prove to be cumbersome. These observations indicate the need for blind adaptive receivers that require no more information than the conventional single-user receivers: the timing and signature waveform of the desired user.

3.4.2.1 LCMV Blind ST MUD and its GSC Implementation

A basic blind adaptive multiuser detection technique based on minimizing the receiver output energy while preserving the desired signal's energy was proposed in [56]. This detection strategy is a special case of a more general optimization technique: the linear constrained minimum variance (LCMV) criterion.

In this subsection we adopt this LCMV criterion to design a blind adaptive space-time multiuser detector. Suppose again the user of interest is the $k$th user. The received signal at the $p$th antenna element is passed through a chip-matched filter and then sampled at the
chip rate to obtain an $\vec{N}$-vector of signal samples $\mathbf{r}_p^{(k)}$ as shown in (3.51). Using (3.5) and (3.52), the signal vector $\mathbf{r}_p^{(k)}$ can be expressed as

$$
\mathbf{r}_p^{(k)} = A_k \mathbf{b}_k \sum_{l=1}^{L} a_{kl,p} g_{kl} s_{kl} + \mathbf{i}_p + \sigma \mathbf{n}_p,
$$

(3.67)

where $\mathbf{n}_p$ is the ambient noise vector, $\mathbf{i}_p$ comprises both MAI and ISI, and the $\vec{N}$-vector $s_{kl}$ is the discretized version of the delayed signature waveform of user $k$ with individual elements given by

$$
s_{kl,n} = \int_{\tau_{kl}}^{\tau_{kl}+(n+1)T_c} s_k(t-\tau_{kl})\psi(t-\tau_{kl1}-nT_c)dt, \quad 1 \leq n \leq \vec{N}, \quad 1 \leq l \leq L.
$$

(3.68)

As in subsection 3.4.1, at the $p$th antenna element, we would like to design a linear filter with coefficients $\mathbf{w}_p^{(k)} = [w_{p,1}^{(k)}, w_{p,2}^{(k)}, \cdots, w_{p,N}^{(k)}]^T$, $1 \leq p \leq P$. But now we trade the knowledge of a training sequence for the desired user for its spreading code. According to the LCMV criterion, we should choose $\mathbf{w}_p^{(k)}$ via the optimization problem

$$
\mathbf{w}_p^{(k)} = \arg \min_{\mathbf{w}_p^{(k)} \in \mathbb{C}^{N}} E\left\{\left\| (\mathbf{w}_p^{(k)})^H \mathbf{r}_p^{(k)} \right\|^2 \right\} = \arg \min_{\mathbf{w}_p^{(k)} \in \mathbb{C}^{N}} (\mathbf{w}_p^{(k)})^H \mathbf{R}_p^{(k)} \mathbf{w}_p^{(k)},
$$

(3.69)

where $\mathbf{R}_p^{(k)} = E[\mathbf{r}_p^{(k)} \mathbf{r}_p^{(k)H}]$ is the autocorrelation matrix of the signal vector $\mathbf{r}_p^{(k)}$, subject to a linear constraint

$$
\mathbf{C}_k^H \mathbf{w}_p^{(k)} = \mathbf{f}_p^{(k)},
$$

(3.70)

where

$$
\mathbf{C}_k = [s_{kl}, s_{k2}, \ldots, s_{kl}].
$$

(3.71)
The constraints (3.70) ~ (3.72) ensure that, after summing the filtered outputs of all antenna elements, the desired signal energy is optimally combined.

The theoretical optimum solution to the above problem (3.69) - (3.70) is given by

$$ (w_p^{(k)})^{opt} = (R_p^{(k)})^{-1} C_k [C_k^H (R_p^{(k)})^{-1} C_k]^{-1} f_p^{(k)}, $$

with the theoretical minimum output energy in (3.69) given by

$$ MOE^{(k)} = (f_p^{(k)})^H [C_k^H (R_p^{(k)})^{-1} C_k]^{-1} f_p^{(k)}. $$

The LCMV criterion has been applied in many signal processing fields, one of which is spatial filtering or beamforming [113]. The generalized sidelobe canceller (GSC) structure for adaptive beamforming represents an effective implementation of the LCMV beamformer, changing a constrained minimization problem into an unconstrained form. We borrow this idea for implementation of our LCMV blind ST MUD, which is shown in Fig. 3.14.

![Diagram](image-url)
The basic idea is to decompose the weight vector \( w^{(k)}_p \) into two orthogonal components: a nonadaptive part \( w^{(k)}_{p,s} \) and an adaptive part \( M^H w^{(k)}_{p,a} \),

\[
\text{w}^{(k)}_p = w^{(k)}_{p,s} - M^H w^{(k)}_{p,a}.
\] (3.75)

The nonadaptive part \( w^{(k)}_{p,s} \) lies in the range of \( C_k \) and fulfills the constraint (3.70), given by

\[
w^{(k)}_{p,s} = C_k (C_k^H C_k)^{-1} C_k^H w^{(k)}_p = C_k (C_k^H C_k)^{-1} f_p.
\] (3.76)

The adaptive part lies in the null space of \( C_k \), whose orthogonality is guaranteed by the \((\tilde{N} - L) \times \tilde{N}\) blocking matrix \( M \), which satisfies

\[
C_k^H M^H = \mathbf{O}_{L \times (\tilde{N} - L)}.
\] (3.77)

There are several ways to obtain \( M \) like Gram-Schmidt orthogonization and QR decomposition. With the constraint of (3.70), the optimization problem has \( \tilde{N} - L \) degrees of freedom. After this decomposition, the constrained minimization problem is transformed to an unconstrained problem for the \( \tilde{N} - L \)-vector \( w^{(k)}_{p,a} \):

\[
w^{(k)}_{p,a} = \arg \min_{w^{(k)}_{p,a}} (w^{(k)}_{p,a} - M^H w^{(k)}_{p,a})^H R_p (w^{(k)}_{p,a} - M^H w^{(k)}_{p,a}),
\] (3.78)

whose theoretical optimum solution is given by

\[
w^{(k)\text{opt}}_p = (M R_p^{(k)} M^H)^{-1} M R_p^{(k)} w^{(k)}_p.
\] (3.79)
Note from Fig. 3.14 that if $(w^{(k)}_{p,s})^H r^{(k)}_p$ is taken as the desired user while $M_r^{(k)}$ is taken as the observations, then the GSC implementation readily lends itself to LMS adaptation with no need of training sequences. The adaptation rule is given as follows:

$$z_p^{(k)}(i) = (w^{(k)}_{p,s} - M_w^{(k)}w_{p,s}(i))^H r^{(k)}_p(i),$$

with

$$w^{(k)}_{p,s}(i+1) = w^{(k)}_{p,s}(i) + \mu z_p^{(k)*}(i)M_r^{(k)}(i).$$

Finally the detected bits are given by

$$\hat{b}_k(i) = \text{sgn}(\text{Re}(z_p^{(k)}(i))).$$

### 3.4.2.2 MIN-MAX Channel Parameter Estimation

The above LCMV ST MUD assumes knowledge of the channel parameters of the desired user (see (3.72)), which in practice should be estimated in advance. This problem can be overcome by using a technique that incorporates the parameter estimation into the LCMV receiver design with a min-max approach, i.e., the idea is to find a constraint vector $f^{(k)}_p$ that maximizes the theoretical minimum output energy (3.74). The problem can be stated as

$$\hat{f}_p^{(k)} = \arg \max_f \text{MOE}^{(k)}(f) \quad s.t. \quad \|f\| = 1$$

where $\text{MOE}^{(k)}(f)$ is defined as in (3.74) in the obvious way. The solution to the above problem is readily given by the minimum-eigenvalue eigenvector of $C^H_k (R_p^{(k)})^{-1} C_k$. It is shown in [106] that if the vectorized signature waveforms of all users and all their
delayed versions are linearly independent, then \( \hat{f}^{(k)}_p \rightarrow f^{(k)}_p \) asymptotically in the sense of norm.

### 3.4.2.3 Robustified Blind ST MUD

In practice, the receiver may assume the original spreading waveform of the desired user as its nominal choice in data detection, whereas the actual received waveform may be distorted during transmission or may include additional unmodeled multipath components. In this situation of signature waveform mismatch, the blocking matrix used in the GSC structure is no longer orthogonal to the desired signal. This lack of orthogonality will lead the adaptive filter \( \mathbf{w}^{(k)}_{p,a} \) to cancel the desired signal in its effort to minimize the output energy. To overcome this problem, we adopt the approach in [56] to constrain the norm of the weight vector to avoid desired signal cancellation; i.e., we introduce a constraint

\[
\| \mathbf{w}_p \| \leq \chi_f, \tag{3.84}
\]

where an approximate choice of \( \chi_f \) can be determined experimentally. The LMS adaptation of this robustified LCMV blind ST MUD is given as follows:

\[
z^{(k)}_p(i) = (\mathbf{w}^{(k)}_{p,r} - \tilde{\mathbf{w}}^{(k)}_{p,a}(i))^H \mathbf{r}^{(k)}_p(i), \tag{3.85}
\]

\[
\mathbf{x}(i) = \tilde{\mathbf{w}}^{(k)}_{p,a}(i) + \Delta z^{(k)*}_p(i) \mathbf{M}^H \mathbf{r}^{(k)}_p(i), \tag{3.86}
\]

and
\[
\tilde{w}_{p, a}^{(k)}(i + 1) = \begin{cases} 
\mathbf{x}(i), & \|\mathbf{x}(i)\| \leq \sqrt{\mathbf{z}_i^2 - \|\mathbf{w}_{p, r}^{(k)}\|^2} \\
\sqrt{\mathbf{z}_i^2 - \|\mathbf{w}_{p, r}^{(k)}\|^2}\frac{\mathbf{x}(i)}{\|\mathbf{x}(i)\|}, & \|\mathbf{x}(i)\| > \sqrt{\mathbf{z}_i^2 - \|\mathbf{w}_{p, r}^{(k)}\|^2}.
\end{cases} 
\] (3.87)

The detected bits are then given by

\[
\hat{b}_k(i) = \text{sgn}(\text{Re}(z_p^{(k)}(i))) .
\] (3.88)

### 3.4.2.4 Numerical Results

In this subsection, the performance of the blind adaptive space-time multiuser detectors is examined through computer simulations. The simulation parameters are similar to those used in subsection 3.4.1.3, with the following exceptions. We assume a \( K = 8 \)-user CDMA system with spreading gain \( N = 16 \). The user of interest is user 1. The delay spread is set to be \( T \). A variable step size is used for the LMS algorithm as follows: 0.01 for the first 250 iterations, 0.005 for the next 250 iterations, and 0.002 afterwards.

Figure 3.15 shows the learning curve for the LCMV blind adaptive ST MUD. The theoretical MOE is also plotted in the figure (as the dashed line) for comparison. Compared to its counterpart in Fig. 3.10, the blind adaptive algorithm converges more slowly and its turbulence around the steady state is more severe, indicating the (expected) inferiority of the performance of the blind adaptive algorithm compared with its data-aided counterpart.

Figure 3.16 validates the claim in subsection 3.4.2.2, i.e., \( \hat{f}_p^{(k)} \to \frac{f_p^{(k)}}{\|f_p^{(k)}\|} \) as the variance of the background noise goes to zero, by showing the normalized error
Fig. 3.15 Convergence of the LCMV blind adaptive space-time multiuser detector

Fig. 3.16 Precision of min/max channel parameter estimation
Fig. 3.17 Performance comparison of LCMV blind adaptive ST MUD with exact and estimated channel parameters

Fig. 3.18 Performance comparison of non-robust and robust LCMV blind adaptive ST MUD in the situation of signature waveform mismatch
\[
\frac{\|f_p^{(k)} - \hat{f}_p^{(k)}\|}{\|f_p^{(k)}\|^2},
\]
where \(\hat{f}_p^{(k)}\) is obtained through (3.83) and \(f_p^{(k)}\) is the actual channel parameter vector for the desired user. Figure 3.17 further compares the performance of the LCMV blind ST MUD with exact and estimated channel parameters. It is seen that min-max channel parameter estimation leads to almost no performance loss.

Finally, Fig. 3.18 shows the robustness of the norm-constrained LCMV blind adaptive ST MUD in the situation of signature waveform mismatch. To simulate the effects of signature distortion, the signature waveform of the desired user is perturbed by Gaussian noise with zero mean and standard deviation 0.07. Note that the performance of the non-robustified LCMV blind adaptive ST MUD worsens as the SNR increases because of the signal cancellation; while that of the robustified LCMV blind adaptive ST MUD greatly outperforms its counterpart, with a 2 dB loss relative to the ideal case without signature waveform mismatch (c.f. Fig. 3.17).

### 3.5 Summary

In this chapter, first we have considered several batch iterative space-time multiuser detection schemes for multipath CDMA channels with multiple receive antennas. Fully exploiting diversities through space-time processing and multiuser detection offers substantial improvement over alternative processing methods. It is shown that iterative implementation of these linear and nonlinear multiuser receivers realizes this substantial gain and approaches the optimum performance with reasonable complexity. Among these iterative implementations the SAGE ST MUD receiver outperforms the others. The
complexity of the new SAGE detector with the traditional structure is no higher than the existing methods but with better performance and smoother convergence. The SAGE detector with the new structure retains its excellent performance but with greater adaptability, and its complexity is comparable to the existing methods. Furthermore, with long (pseudo-random) spreading codes (e.g. IS-95 and its 3G counterparts) or in the rapidly time-varying environments, the SAGE detector with the new structure exhibits some advantage as it circumvents the system matrix update problem.

Another topic considered in this chapter is sample-by-sample adaptive space-time multiuser detectors. The data-aided adaptive space-time multiuser receivers combine the functions of adaptive beamforming, RAKE combining and multiuser detection with no side information needed other than the timing and training sequences of the desired user. The alternative blind adaptive space-time multiuser receiver is based on the LCMV criterion and min-max parameter estimation. This detector is robustified against signature waveform mismatch with norm-constrained techniques. LMS implementations of all of these adaptive ST MUD receivers have been considered, and the convergence and excess-MSE issues have been discussed. Several issues associated with the algorithms deserve further study. For example, faster adaptive algorithms such as RLS can be considered. Also efforts might be made to narrow the performance gap between blind adaptive techniques and data aided adaptive techniques. For example, techniques developed in the context of blind source separation or blind equalization, like fourth-order cumulant-based methods and the constant modulus algorithm (CMA), can be considered to apply to the blind space-time multiuser detection problem studied in this chapter.
MIMO Systems: Turbo Space-Time Multiuser Detection

4.1 Introduction

Recent information theoretic results have indicated the enormous capacity potential of wireless communication systems with antenna arrays at both the transmitters and receivers. These so-called multiple-input multiple-output (MIMO) systems have been shown to yield a tremendous capacity, which grows at least linearly with the minimum of the numbers of transmit and receive antennas [42], [105].

The enormous spectral efficiencies of MIMO systems were obtained for a single link with white Gaussian noise. In a cellular environment, there will often be co-channel interference from other cells, which becomes the dominating channel impairment. It was shown in [15] that in an interference-limited environment, the capacity of a MIMO system is hardly larger than when using smart antennas at the receivers only. This seems to be related to the fact that an antenna array with $N$ elements can eliminate $N-1$ interferers, so that the reuse distance (in a TDMA/FDMA system) can be chosen to be very small. The independent data streams employed by a MIMO system are all different interferers, so a receive array has no degrees of freedom with which to cancel the co-
channel interferers after it separates the multiple data streams in its own cell. On the other hand, this investigation assumed a certain system structure taken from the noise-limited case, and did not try to optimize the system for interference-limited environments. To be specific, they exploited sub-optimum signal processing techniques (uncoded V-BLAST) at the receivers; no attempt was made to jointly detect desired as well as interfering signals; and no cooperation between base stations was assumed.

Our study now investigates whether a more advanced receiver structure can significantly increase the capacity of MIMO-systems with adjacent-cell interference. Any BLAST-like (Bell Labs layered space-time architecture) scheme is by its nature a multiuser detector that separates the data streams from the antennas of the desired base station. It thus seems logical to extend that principle also to the data streams from the interfering base stations. In this chapter, turbo space-time multiuser detection (ST MUD) is employed for intracell communications, then on top of it various multiuser detection methods are applied to combat intercell interference, thereby hopefully to increase the capacity in this interference-limited scenario. We concentrate here on the downlink, as this is usually the bottleneck for wireless data transmission. Furthermore, we assume that there is no cooperation between base stations (e.g., no joint transmission as in [6], [97]), and that the base stations have no knowledge of the downlink propagation channel. These assumptions are well fulfilled in typical wireless LAN situations. In the end, however, we will address whether it is worth to request more system resource for performance improvement.

This chapter is organized in the following way: in Section 4.2, the system model and the assumptions made in the problem formulation are presented. In Section 4.3, turbo
space-time multiuser detector structure for intracell communications is illustrated. In Section 4.4, various potential multiuser detection methods are introduced to combat the intercell interference. Next, in Section 4.5 various multiuser detection schemes are examined; and an adaptive detection scheme is proposed, which together with an advanced turbo ST MUD structure offers substantial performance gain over the well-known V-BLAST techniques with coding in this interference-limited cellular environment. We also compare our results to single-cell upper capacity bounds, and show that significant gains can be made by base-station cooperation algorithms. Conclusions and some insights are given in Section 4.6.

4.2 Problem Formulation

4.2.1 MIMO System Model

For the single-cell interference-free case, Teletar [105] and Foschini [42] have derived the exact capacity expressions as well as useful approximations and lower bounds for MIMO systems. We adopt the same mathematical model as in [105] and [42], which is given by

$$y = Hx + n,$$  \hspace{1cm} (4.1)

where $y$ is the received vector, $x$ is the transmitted signal, $H$ is the channel matrix which captures the channel characteristics between transmit and receive antenna arrays, and $n$ is the background noise. Without loss of generality, we assume an $N \times N$ MIMO system with the transmitted signal vector constrained to have overall power $E[x^H x] \leq P$, and circularly symmetric Gaussian background noise with covariance matrix $\Phi_n = \sigma^2 I$. The
entries of the complex matrix $H$ are independent with uniformly distributed phase and normalized Rayleigh distributed magnitude, modeling a Rayleigh fading channel with sufficient physical separation between transmit and receive antennas. The signal-to-noise ratio (SNR) is given by $\rho = P/\sigma^2$. We list the following formulas from [105] and [42] below that will serve as the basis for their extension to multi-cell interference-limited cases. In all cases, the channel state information (CSI) is assumed known at the receiver.

If the channel matrix $H$ is known at the transmitter, then the capacity is given by

$$C_U = \sum_i \log_2(1 + \lambda_i P_i), \quad (4.2)$$

where $\{\lambda_i\}$ are eigenvalues of the matrix $H^H\Phi^{-1}H$, and the values of $P_i$, i.e., the powers assigned to the different eigenmodes, are derived from the water-filling rule

$$P_i = (\theta - \frac{1}{\lambda_i})^+, \quad (4.3)$$

where $x^+ = \max(x, 0)$ and $\theta$ is chosen to satisfy $\sum P_i = P$.

If the channel matrix $H$ is unknown at the transmitter (as is assumed in our study), then the capacity is given by

$$C = \log_2 \det[I + \frac{P}{N} H^H\Phi^{-1}H], \quad (4.4)$$

which can be viewed as an equal power allocation to the eigenmodes due to the lack of channel knowledge.

If the channel matrix $H$ is unknown at the transmitter, a lower bound on the capacity with $\Phi_N = \sigma^2 I$ is given by
\[ C_L = \sum_i \log_2(1 + \frac{P}{N^2 \chi^2_{2i}}), \quad (4.5) \]

where \( \chi^2_{2i} \) is a \( \chi^2 \) distributed random variable with \( 2i \) degrees of freedom and mean value \( i \).

### 4.2.2 Cellular System Model

We consider a TDMA/FDMA multi-cell system (or equivalently an orthogonal CDMA system), where each base station (BS) and mobile station (MS) has the same number of antennas \( N \). We take into account interference from the first tier of the center-excited cell configuration with reuse factor of one, which is depicted in Fig. 4.1. Note that we mainly deal with the wireless LAN application with pico-cells, so no sectorization of the cell is intended. We assume a frequency-flat, quasi-static fading environment, and the complex baseband channel gain between the \( j \)th transmit and the \( i \)th receive antenna is modeled by

\[ h_{ij} = \sqrt{\frac{1}{d_{ij}}} \sqrt{s_{ij}} \left[ \sqrt{\frac{K}{K+1}} e^{j \phi_j} + \sqrt{\frac{1}{K+1}} z_j \right], \quad (4.6) \]

where the three terms embody the path loss, the shadow fading and multipath fading effect, respectively. In particular, we have the following parameters.

**Path loss**: \( d_{ij} \) is the length of the link and \( \gamma \) is the path loss exponent;

**Shadow fading**: \( s_{ij} = 10 \log_{10} S_{ij} \) is a log-normal shadow fading variable, where \( S_{ij} \) is a zero mean Gaussian random variable with standard deviation \( \nu \);

---

\(^1\) The sum of the squares of \( 2i \) real Gaussian variables, each with zero mean and variance 0.5.
Multipath fading: $K$ is the so-called Ricean $K$-factor, which denotes the ratio of the direct received power (line-of-sight (LOS) component) to average scattered power (non-LOS (NLOS) component); $\Phi_{ij} = \frac{2\pi d_y}{\lambda}$ is the phase shift of the LOS path ($\lambda$ is the wavelength); $z_y$ is modeled as a set of normalized complex Gaussian random variables, assumed to be independent for each transmit-receive link.

With these assumptions, the multicell system model is given by

$$ y = H \cdot x + \sum_i H_{ij} \cdot x_{ij} + n, $$

(4.7)

where the subscript “ij” denotes interference. The channel matrices $H$ and $\{H_{ij}\}$ are independent with independent and identically distributed (i.i.d.) elements given by (4.6). The transmitted signals from all users are assumed to be of the same format with
$E[x^H x] = E[x_{gi}^H x_{gi}] \leq P$, whose codebooks are known to the receivers. The noise is assumed to be white and complex Gaussian with covariance matrix $\Phi_n = \sigma^2 I$.

In order to make the analysis more tractable, the multicell scenario is usually simplified to a linear array of cells and the interference from the two adjacent cells is characterized by a single attenuation factor [129]. To provide a common framework that is general enough to address multiuser detection across the cell while remaining simple enough for analysis and simulation, we simplify our model so that there are four interferers in two groups of two, in which one group is much stronger than the other. This roughly reflects the essential reality as interference from two farthest adjacent cells can typically be ignored, and simulation results verify that the power of the two strongest users usually dominates. Thus, the model (4.7) is modified as

$$y = H \cdot x + \sum_{j=1}^{2} H_{g_j} \cdot x_{g_j} + \sum_{i=3}^{4} H_{g_i} \cdot x_{g_i} + n,$$

(4.8)

with $P_{g_1} = \alpha P_{g_2}$, $P_{g_3} = \beta P_{g_4}$, and $\frac{(P_{g_1} + P_{g_2})}{(P_{g_3} + P_{g_4})} = \gamma >> 1$, where $P_{g_i} = E[x_{gi}^H x_{gi}]$. The different choice of the parameters $\alpha$, $\beta$, and $\gamma$ defines the structure of the interfering signals, to be further addressed in Section 4.5. We use the same assumptions for the channel matrices and noise as (4.1), while assuming the channel matrices for different cells are independent. The signal-to-noise ratio is given by $\rho = P/\sigma^2$, and the signal-to-interference ratio (SIR) is given by $\eta = \frac{P}{\sum_i P_{g_i}}$. 

We will mainly use model (4.8) for our study. In the end, however, results with model (4.7) will also be given to test and validate the proposed algorithms with more realistic settings.

4.3 Turbo Space-Time Multiuser Detection for Intracell Communications

In this section, let us assume a single cell scenario for ease of illustration. We will address the multicell case in the next section.

4.3.1 Receiver Structures and Diversity

References [41] and [43] propose two layered space-time architectures, called D-BLAST, and V-BLAST, respectively. Actually, the space-time layered architecture falls into the larger category of space-time multiuser detection, which refers to the application of the multiuser detection techniques with the aid of both temporal (e.g. CDMA codes) and spatial (spatial signature) structures of the signals to be detected. The BLAST technique is essentially a decision feedback space-time multiuser detector.

In recent years, iterative processing techniques with soft-in/soft-out (SISO) components have received considerable attention. The basic idea is to break up optimum joint signal processing, e.g. concatenated decoding, joint equalization and decoding, or joint decoding and multiuser detection, into simpler separate components, iterating between them with the exchange of probabilities or “soft” information. This approach typically performs almost as well as the much more complex “joint” approach which
attempts to achieve the exact ML or MAP optimization. This so-called turbo principle is exemplified through turbo decoding [51], turbo equalization [32] and turbo multiuser detection [79] with application to wireless [126] and wireline communications [139].

The turbo multiuser detection can be applied to the coded BLAST system, resulting in two turbo space-time multiuser detection structures, shown in Fig. 4.2 and Fig. 4.3, respectively. One is called coded V-BLAST, where at the transmitter the information bits are first demultiplexed into $N$ substreams, each of which is independently encoded, interleaved, and symbol-mapped. At the receiver, the MMSE criterion is used to decouple the substreams; then for each substream a soft metric is calculated and fed to the SISO MAP decoder, which produces soft estimates of information and coded bits, used to refine soft metric calculation in the next iteration. After several iterations within a layer, the estimated bits are good enough to be used as output as well as to be fed to the next layer to assist in detection. The other is called Turbo-BLAST, where at the transmitter the information bits are coded (not necessarily with turbo codes) and interleaved as a whole; then the whole coded stream is demultiplexed into $N$ substreams and symbol-mapped individually. At the receiver, the entire data stream is processed iteratively between a soft metric calculation stage and a decoding stage. Note that in the soft metric calculation stage, either a maximum likelihood (ML) joint detection or a MMSE multistage parallel interference cancellation (PIC) scheme can be used. We will show that these two schemes achieve the same performance, owing to the turbo processing.

For the coded V-BLAST, each substream is tied to a fixed antenna element so no transmit diversity is exploited. On the contrary, Turbo-BLAST, like D-BLAST, introduces inter-substream coding and takes advantage of transmit diversity with transmit
antenna arrays. At the receiver end, the first detected substream of the V-BLAST will essentially determine the overall system performance due to error propagation. Unfortunately, it has the least receive diversity degree as a result of interference cancellation. This is also true for D-BLAST. However, for the Turbo-BLAST, either ML MUD or the less-complex MMSE PIC brings in full receive diversity. Therefore, the Turbo-BLAST is expected to even outperform the coded D-BLAST, which can achieve a tight lower bound (4.5) on the capacity theoretically. In Section 4.5, it is shown that the Turbo-BLAST structure essentially approaches the capacity (4.4) in the interference-free case. The V-BLAST structure serves mainly as a baseline in this study, as it is the first implemented space-time layered architecture and the most promising one to be employed in commercial wireless LAN application, due to its simplicity. (The study of D-BLAST is mainly in the information-theoretic aspect.)

![Fig. 4.2 Structure of coded V-BLAST](image)
4.3.2 Turbo-BLAST Detection

The turbo decoding procedure of the coded V-BLAST is exactly analogous to that of the Turbo-BLAST to be discussed and therefore is omitted here. The Turbo-BLAST detection algorithm involves two components: demodulation and decoding. The demodulation stage with ML is straightforward. Suppose an $N \times N$ MIMO system is employed by one cell, and each substream adopts $M$-QAM (quadrature amplitude modulation). Then for each symbol interval $B = N \cdot \log_2 M$ bits are jointly detected. The extrinsic information for the $k$th bit of the $i$th substream symbol, $1 \leq k \leq \log_2 M$, $1 \leq i \leq N$, is given by

$$
\hat{\lambda}_i(b_{k,i}) = \Lambda_i(b_{k,i}) - \hat{\lambda}_i^p(b_{k,i})
$$

$$
= \log \frac{\sum_{x \in \Lambda_i} p(y \mid x)p(x)}{\sum_{x \in \Lambda_i} p(y \mid x)p(x)} - \log \frac{P(b_{k,i} = 1)}{P(b_{k,i} = -1)},
$$

(4.9)
where \( X^+_i = \{(x_1, x_2, \ldots, x_N)^T : b_{k,j} = 1\} \) and \( X^-_i = \{(x_1, x_2, \ldots, x_N)^T : b_{k,j} = -1\} \); \( p(y \mid x) \) is a multivariate Gaussian distribution (see (4.1)); \( \lambda^*_2(b_{k,j}) = \log\left(\frac{P(b_{k,j} = 1)}{P(b_{k,j} = -1)}\right) \) and \( p(x) = \prod_{k,j} p(b_{k,j}) \) comprise a priori information from the decoding stage.

The demodulation stage with PIC is subtler. Suppose the received signal for some substream \( 1 \leq k \leq N \) after interference cancellation is given by

\[
\tilde{y}_k = H(x - \tilde{x}_k) + n, \tag{4.10}
\]

where \( \tilde{x}_k = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k = 0, \ldots, \tilde{x}_N)^T \) is the estimated interference vector. First an MMSE filter is applied to \( \tilde{y}_k \) to further suppress the residual interference plus noise, that is,

\[
w_k = E[\tilde{y}_k \tilde{x}_k^H]^{-1} E[\tilde{y}_k x_k^*] = (h_k^H + H_k Q H_k^H + \frac{N}{\rho} I)^{-1} h_k, \tag{4.11}
\]

where \( h_k \) is the \( k \)th column of matrix \( H \), \( H_k \) is the complement of \( h_k \) in \( H \), and

\[
Q = \text{diag}[1 - \frac{\nu}{N} \mid x_1 \mid^2, \ldots, 1 - \frac{\nu}{N} \mid x_{k-1} \mid^2, 1 - \frac{\nu}{N} \mid x_{k+1} \mid^2, \ldots, 1 - \frac{\nu}{N} \mid x_N \mid^2],
\]

which approaches 0 when estimates from the decoding stage are accurate enough for constant-modulus signals. As is shown in [80], the output of the MMSE filter \( z_k = w_k^H \tilde{y}_k \) can be written as

\[
z_k = \mu_k x_k + \eta_k, \tag{4.12}
\]

where \( \mu_k = \frac{\nu}{N} E[z_k x_k^*] = w_k^H h_k \), and \( \eta_k \) is a Gaussian variable with zero mean and variance \( \nu_k^2 = E[|z_k - \mu_k x_k|^2] = E[|z_k|^2] - \frac{\nu}{N} |\mu_k|^2 = \frac{\rho}{\nu} (\mu_k - |\mu_k|^2) \). The extrinsic information is given in the same form as (4.9), but with \( y \) replaced by \( z_k \) and \( x \) with \( x_k \), and (4.1) replaced with (4.12), and therefore with much lower complexity.
For the channel decoding stage let us consider a binary $1/n$ convolutional code with constraint length $\nu$ (the soft decoding for other codes, e.g., Turbo codes, is similar). The BCJR decoding algorithm [5] is known to yield the optimal symbol estimate with minimum symbol-error-rate. What is more important for turbo processing is the BCJR algorithm’s ability to yield soft information in the form of \textit{a posteriori} log-likelihood ratios (LLRs) for coded and information bits. Using the same notation as in [5], for stage $t$ of the code trellis transiting from state $S_{t-1} = s'$ to $S_t = s$ associated with input $d_t$ and output $b_t = (b_t^1, ..., b_t^n)$ (where $\{b_t^j\} \leftrightarrow \{b_t^k\}$ with $j = (t-1)n + k$ for a rate-$1/n$ convolutional code), we have

\[
\Lambda(b_t^k) = \log \frac{\sum_{s_t^k} \alpha_{t-1}(s') \gamma_t(s', s) \beta_t(s)}{\sum_{s_t} \alpha_{t-1}(s') \gamma_t(s', s) \beta_t(s)},
\]

(4.13)

where $S^+_k$ is the set of state pairs $(s', s)$ such that the $k$th coded bit at stage $t$ is 1 and $S^-_k$ is the corresponding set for -1; and

\[
\Lambda(d_t^j) = \log \frac{\sum_{u_t^j} \alpha_{t-1}(s') \gamma_t(s', s) \beta_t(s)}{\sum_{u_t} \alpha_{t-1}(s') \gamma_t(s', s) \beta_t(s)},
\]

(4.14)

where $U^+_k$ is the set of state pairs $(s', s)$ such that the information bit at stage $t$ is 1 and $U^-_k$ is the corresponding set for -1. The $\alpha_t(s)$, $\beta_t(s)$ and $\gamma_t(s', s)$ terms are defined as follows. $\gamma_t(s', s)$ denotes the transition probability for the branch $s' \rightarrow s$ for the stage $t$ of the code trellis. As no channel outputs are available for the outer code (the convolutional
code is treated as the outer code while the modulation is treated as the inner code for this concatenated system) of a concatenated coding system, we have

\[ \gamma_i(s', s) = P(S_i = s \mid S_{i-1} = s') = \prod_{l=1}^{n} P(b^l_j) = \prod_{l=1}^{n} \frac{1}{2} [1 + b_j \tanh(\lambda^p_l(b_j))], \quad (4.15) \]

where \( \{b_j\} \leftrightarrow \{b^k_j\} \) with \( j = (t-1)n + k \), and \( \lambda^p_l(b_j) \) represents the corresponding a priori LLR delivered from the demodulation stage that is related to the state transition of \( s' \rightarrow s \). The \( \alpha_t \) and \( \beta_t \) terms are defined with forward and backward recursions as

\[
\alpha_t(s) = \sum_{s'} \alpha_{t-1}(s') \gamma_t(s', s), \quad t = 1, 2, \ldots, \tau, \quad (4.16)
\]

and

\[
\beta_t(s) = \sum_{s'} \beta_{t+1}(s') \gamma_{t+1}(s, s'), \quad t = \tau - 1, \tau - 2, \ldots, 0, \quad (4.17)
\]

with boundary conditions

\[
\alpha_0(0) = 1, \quad \text{and} \quad \alpha_0(s) = 0, \quad \text{for} \quad s \neq 0, \quad (4.18)
\]

and

\[
\beta_\tau(0) = 1, \quad \text{and} \quad \beta_\tau(s) = 0, \quad \text{for} \quad s \neq 0, \quad (4.19)
\]

where \( \tau \) denotes the frame length of the information bits. The summations of (4.16) and (4.17) are over all states \( s' \) for which the transition \( s' \leftrightarrow s \) is possible. So the extrinsic information produced by the SISO channel decoder can be written as

\[
\lambda_2(b^k_j) = \log \frac{\sum_{S_k} \alpha_{t-1}(s') \beta_t(s) \prod_{l=1,l \neq k}^{n} P(b^l_j)}{\sum_{S_k} \alpha_{t-1}(s') \beta_t(s) \prod_{l=1,l \neq k}^{n} P(b^l_j)} = \Lambda_2(b^k_j) - \lambda^p_l(b^k_j), \quad (4.20)
\]
where again \( \{b_j\} \leftrightarrow \{b_k^\dagger\} \) with \( j = (t-1)n + k \) for a rate-1/\( n \) convolutional code.

The BCJR algorithm is known to have numerical problems associated with the representations of probabilities due to the large dynamic range of \( \alpha_t \) and \( \beta_t \). Thus it is preferable that operations be processed in the logarithmic domain [90], [123]. With the substitution of \( a_i = \log(\alpha_i), \ b_i = \log(\beta_i) \) and \( c_i = \log(\gamma_i) \), (4.13) can be rewritten as

\[
\Lambda(b_t^\dagger) = \log\left( \sum_{S_i} \exp(a_{t-1} + c_i + b_i) \right) - \log\left( \sum_{S_i} \exp(a_{t-1} + c_i + b_i) \right), \tag{4.21}
\]

where

\[
c_i(s', s) = \sum_{l=1}^{n} \log(P(b_l^\prime)) = \sum_{l=1}^{n} \log(\frac{\exp(b_l^\prime \lambda_s^{\prime l}(b_l^\prime))}{1 + \exp(b_l^\prime \lambda_s^{\prime l}(b_l^\prime))}), \tag{4.22}
\]

\[
a_i(s) = \log\left( \sum_{s'} \exp(a_{i-1}(s') + c_i(s', s)) \right), \tag{4.23}
\]

with

\[
a_0(0) = 0, \quad \text{and} \quad a_i(s) = -\infty, \quad \text{for} \quad s \neq 0, \tag{4.24}
\]

and

\[
b_i(s) = \log\left( \sum_{s'} \exp(b_{i-1}(s') + c_{i-1}(s, s')) \right), \tag{4.25}
\]

with

\[
b_0(0) = 0, \quad \text{and} \quad b_i(s) = -\infty, \quad \text{for} \quad s \neq 0. \tag{4.26}
\]
Using the approximation

\[
\log \sum_i e^{x_i} \approx \max_j x_j, \tag{4.27}
\]

Equations (4.21), (4.23) and (4.25) become

\[
\lambda(b^k_i) = \max_{S_i} \left[ (a_{i-1} + c_i + b_i) \right] - \max_{S_i} \left[ (a_{i-1} + c_i) \right], \tag{4.28}
\]

\[
a_i(s) = \max_{s'} \left[ a_{i-1}(s') + c_i(s', s) \right], \tag{4.29}
\]

and

\[
b_i(s) = \max_{s'} \left[ b_{i-1}(s') + c_i(s, s') \right], \tag{4.30}
\]

where (4.29) and (4.30) can be readily recognized as forward and backward Viterbi algorithms. This simplified maximum \textit{a posteriori} probability (MAP) algorithm is the Max-Log-MAP algorithm.

Another approximation of the MAP algorithm is the soft-output Viterbi algorithm (SOVA) [50]. In traditional convolutional decoding, the Viterbi algorithm (VA) is known to be optimal for bit sequence detection and to perform almost as well as the MAP algorithm but with much lower complexity. One of the drawbacks of the VA is its inability to produce soft output information, which hurts the performance of systems in a concatenated form. The SOVA corrects this problem by not only finding the most likely path sequence in a finite state Markov-chain but also delivering additionally a reliability value for each coded and information bit. The fundamental information for allocating a reliability value to the choice of one path, the survivor, of the two paths which merge in a node of a binary code trellis, is the difference between the two accumulated metrics associated with this node. This reliability information is set as
\[ \Delta_i = M_2(t) - M_1(t) = \max_{s'} \left[ a_{t-i}(s') + c_i(s', s) \right] - \min_{s'} \left[ a_{t-i}(s') + c_i(s', s) \right], \quad (4.31) \]

for some node \( s \) of the maximum likelihood sequence at time \( t \), where \( M_2(t) \) denotes the accumulated metric to time \( t \) of the survivor path while \( M_1(t) \) denotes that of the discarded path. The probability that we make the correct choice is given by

\[ P(\text{correct}) = \frac{e^{M_2}}{e^{M_1} + e^{M_2}} = \frac{e^{\Delta_i}}{1 + e^{\Delta_i}}. \quad (4.32) \]

So,

\[ \Delta_i = \log \frac{P(\text{correct})}{1 - P(\text{correct})} \quad (4.33) \]

is just the LLR for this binary path decision. Suppose we want to obtain soft output for a coded bit \( b^k_t \), which the SOVA decides after a delay of \( \delta \). Then on the section of the maximum likelihood path from stage \( t \) to \( t + \delta \), \( \delta + 1 \) nonsurviving paths \( s^a_t, \ldots, s^a_{t+\delta} \) have been discarded. Let us denote the soft information along this section of the ML path by \( \Delta_i, \ldots, \Delta_{i+\delta} \), and examine for each nonsurviving path whether \( \hat{b}^k_t \neq (\tilde{b}^k_t)^a_{t+l} \), where \( \hat{b}^k_t \) is the ML estimate, and \( (\tilde{b}^k_t)^a_{t+l} \) is the associated \( k \)th bit at time \( t \) for the nonsurviving path \( s^a_{t+l}, 0 \leq l \leq \delta \). Then the SOVA-based LLR for \( b^k_t \) is approximately given by

\[ \Lambda_2(b^k_t) \cong \hat{b}^k_t \cdot \min_{0 \leq l \leq \delta} \Delta_{t+l}, \quad (4.34) \]

where the minimum is taken over only those nonsurviving paths which would have a different estimate for \( b^k_t \). Similar results can be obtained for the LLRs of the information bits. Note that we have assumed the ML path has already been obtained before tracing
back to decide the LLR value. Modification of this procedure for truncated ML path selection is straightforward.

Now that we have described the three most commonly used SISO decoding algorithms, we would like to make a brief comparison between them. The MAP algorithm takes account of all the paths in the trellis and divides them into two sets (see (124)), corresponding to the specific binary value of the bit of interest. This division varies from stage by stage. In contrast, the Max-Log-MAP considers only two paths, the ML path and another closest path to the ML path differing only at the transition associated with the bit of interest (see (4.28)). Thus, Max-Log-MAP is suboptimal with respect to MAP with simplified metric computation. However, the loss due to this suboptimality is typically within 0.5 dB for AWGN channels and is negligible for realistic channels, because the Max-Log-MAP algorithm is more robust when the system has estimation errors or the noise is not strictly Gaussian [68]. Furthermore, by replacing logarithms by Jacobian logarithms, defined as

\[
\log(e^{x_1} + e^{x_2}) = \max(x_1, x_2) + \log(1 + e^{\max(x_2 - x_1)}), \tag{4.35}
\]

Max-Log-MAP can be transformed back to exact MAP. Compared to (25), this change requires the addition of only one correction term, which can easily be implemented with a one-dimensional lookup table [90]. Compared to Max-Log-MAP, SOVA also deals with only two paths: one is the ML path, but the other is not necessarily the same as that considered in Max-Log-MAP, the best competing path. Thus, SOVA will lose approximately another 0.5 dB in performance compared with Max-Log-MAP, but with a much lower complexity [68]. Again, it is shown in [44] that some modifications can be made on SOVA to make it equivalent to Max-Log-MAP.
4.4 Multiuser Detection to Combat Intercell Interference

We have already discussed various MUD schemes for detection of different substreams within a MIMO system (intracell interference). Here we will focus on exploiting MUD to combat interference of the same format from adjacent cells (intercell interference).

4.4.1 Maximum Likelihood MUD

Maximum likelihood multiuser detection is infeasible for most current applications due to its complexity. Suppose an \( N \times N \) MIMO system is employed by one cell, and each substream adopts \( M \)-QAM. Then the ML-MUD complexity would be in the order of \( M^N \). If we want to jointly detect all the information bits for users from desired and \( K-1 \) interfering cells, then the complexity would go to the order of \( M^{NK} \). Even if we assume the simplest scheme such as \( M = 4 \), \( N = 2 \) and \( K = 5 \) (ignoring the two weakest interfering cells of the first tier), the complexity would be in the order of \( 2^{20} \), which is beyond the capacity of all current practical systems.

4.4.2 Linear MMSE MUD

We assume knowledge of channel information for the interfering users, which can be obtained either through a joint training phase with the coordination of base stations, or through adaptive tracking algorithms from the received signals directly. MMSE MUD, which is generally the most favorable linear MUD, has a detection matrix given by (c.f. (4.8))

\[
W = \left( HH^H + \sum_i \frac{P_{gi}}{P} H_{gi} H_{gi}^H + \frac{N}{\rho} \right)^{-1} H. \tag{4.36}
\]
Thus, the detection process would be to first apply the weight matrix of (4.36) to the received signal (4.8) to combat CCI; and then to process the modified signal as in Section 4.3. As we mentioned, linear MMSE MUD cannot effectively suppress the intercell interference as the receive antenna array does not have enough degrees of freedom. However, the distribution of the residual interference plus noise at the output of a linear MMSE multiuser detector is well approximated by a Gaussian distribution [80]. This property will guarantee good performance of the Gaussian-metric-based receivers, which would otherwise deteriorate greatly in a multiuser environment.

4.4.3 Linear Channel Shortening MUD

Another linear MUD technique of interest to combat the intercell interference is the so-called channel-shortening multiuser detector, which employs a slightly different optimization criterion than linear MMSE MUD above [70]. For detecting data originating in the desired cell, the idea is to apply some form of array processing to maximize the SINR, where the signal power refers to the power contributions of all the users in the cell to be detected, while interference refers to the power contributions of users in other cells. Note that this criterion is different from linear MMSE MUD (which also maximizes the SINR) in which the signal refers to the very user to be detected while all other users both in cell and out of cell are treated as interferers. In short, the optimal detection matrix for channel-shortening linear MUD is the collection of the first \( N \) principal general eigenvectors of the matrix pencil \( \left( HH^H, \sum_i \frac{P_i}{P} H_{i\ell} H_{i\ell}^H + \frac{N}{\rho} I \right) \). This scheme also serves as a linear pre-processing stage, often followed by much more complex processing, such as ML processing, within the desired cell.
4.4.4 Group IC MUD

Since ML-MUD is highly complex, while linear MUD is limited in its interference cancellation capability (especially for heavy-loaded systems), non-linear MUD often provides a tradeoff between performance and complexity, as noted in Chapter 3. In the context of multicell MIMO systems, group detection techniques naturally call for attention, in which information bits for one group (one cell MIMO) are detected at a time. Following a natural extension from BLAST, we can detect one MIMO system at a time, and feed decisions to other group detectors for interference cancellation. Successive interference cancellation, even though far from the optimum detection scheme, is nonetheless asymptotically optimal under the assumption of perfect interference cancellation [117]. Note that in practice, the success of interference cancellation relies on the correct detection of interference. In adverse environments where we cannot get good estimates of interference, IC schemes will worsen the performance instead of improving it. The potential benefit of group IC MUD depends highly on the interference structure, which will be addressed further in the next section.

4.5 Numerical Results

4.5.1 Comparison of Various MUD Schemes for Intercell Interference Mitigation

In Section 4.3 and 4.4, various potential advanced techniques have been introduced, the combination of which could yield many detector structures. We now compare them, based on the model (4.8), to see which one performs best in the interference-limited
environment. The performance measure is the block-error rate (BLER) over Rayleigh fading channels.

Before conducting simulations, we investigate the distribution of the interference signal strength in a typical scenario. To this end, we set up a simulation scenario for the downlink cellular system with one tier of interferers as shown in Fig. 4.1. We assume a center-excited pico-cell structure with radius $d = 200$ m. The transmit antenna array sends out signals simultaneously from all elements with a total power of 1W at the 2.45 GHz band, which undergo free-space path loss up to a distance of 10m, and then suffers path loss according to a power law with exponent $\eta = 3.7$. The lognormal shadow fading standard deviation $\nu = 8$ dB and Ricean $K$-factor $= 0$. The multipath fading is assumed to be zero-mean complex Gaussian with variance 0.5 per dimension. Each user is randomly located, according to a uniform distribution over the cell. The cumulative distribution functions (CDF) of SNR and SIR that a mobile station experiences are shown in Fig. 4.4 and Fig. 4.5, respectively. The 90th percentile of SNR is 27 dB while that of SIR is 0 dB, which clearly indicates that the environment is interference-limited. Fig. 4.6 indicates that in most cases the power of the two strongest users dominates. A somewhat surprising phenomenon is shown in Fig. 4.7, which indicates that the one-dominant-interferer scenario (the power of the strongest interferer is at least 3 dB higher than the sum of rest) accounts for one third of all the cases. We also found that for the rest two-thirds cases, which belongs to the two-dominant-interferer scenario as indicated by Fig. 4.6, the ratio between the two largest interferer powers varies mostly from 0 – 5 dB.
Fig. 4.4 CDF of SNR experienced by a mobile

Fig. 4.5 CDF of SIR experienced by a mobile
Fig. 4.6 CDF of the ratio between the power sum of the two strongest interferers and the power sum of the remaining interferers experienced by a mobile

Fig. 4.7 CDF of the ratio between the power of the strongest interferer and the power sum of the remaining interferers experienced by a mobile
We assume that each cell employs a 4×4 MIMO system, operating at SNR = 30 dB. The modulation scheme employed is 4QAM. The coding scheme used is a rate-1/3 64-state convolutional code with generators $(G_1, G_2, G_3) = (155,117,123)_8$ (as proposed for EDGE). It was shown in our simulations that this code achieves better performance than a turbo-code with two identical 16-state recursive encoders with generators $(G_1, G_2) = (23,31)_8$, at a considerably lower complexity. We transmit blocks of 384 information bits, and record the block error probability of this system.

The receiver structure is either V-BLAST or Turbo-BLAST, combined with various MUD schemes. To be specific, the receivers we study are: 1) Coded V-BLAST (V-BLAST); 2) Coded V-BLAST with linear MMSE MUD pre-processing (V-BLAST+MMSE); 3) Turbo-BLAST with a parallel interference cancellation demodulation stage (T-BLAST (PIC)); 4) Turbo-BLAST with a parallel interference cancellation demodulation stage, with linear MMSE MUD pre-processing (T-BLAST (PIC)+MMSE); 5) Turbo-BLAST with a maximum likelihood demodulation stage (T-BLAST (ML)); 6) Turbo-BLAST with a maximum likelihood demodulation stage, with linear channel shortening MUD pre-processing (T-BLAST (ML)+CS); 7) Turbo-BLAST with a parallel interference cancellation demodulation stage, with full group IC MUD (T-BLAST (PIC)+IC). We study the performance of these receivers in the framework of (4.8) in two situations: (A) $P_{y_1} = P_{y_2} = 4P_{y_3} = 4P_{y_4}$ and (B) $P_{y_1} = 6P_{y_2} = 6P_{y_3} = 6P_{y_4}$. Situation (A) corresponds to a two-equal-power-dominant-interferer scenario, while situation (B) reflects a one-dominant-interferer case.

---

2 This receiver attempts to detect all the interfering signals of interest
The simulation results for situation (A) are shown in Fig. 4.8, from which we can see that: 1) Turbo-BLAST offers both diversity\(^3\) and coding\(^4\) gain over V-BLAST; 2) Turbo-BLAST with a PIC demodulation stage performs as well as Turbo-BLAST with ML stage, while it has much lower complexity; 3) Linear MUD pre-processing offers a considerable performance gain in interference-limited environments; and 4) Group IC MUD worsens the performance due to incorrect decision feedback. Note that we attempt to detect all interfering signals in this case. In all, we see that Turbo-BLAST with linear MMSE MUD to combat the intercell interference achieves the best performance, which is about 2 dB and 6 dB over Turbo-BLAST and coded V-BLAST, without MUD, respectively, at 1% BLER.

![Fig. 4.8 Performance comparison of various MIMO receivers when two equal-power interferers dominate](image)

\(^3\) The slope of BLER curves
\(^4\) The horizontal distance between BLER curves
The failure of group IC MUD is owing to the inability to correctly detect the information bits for interfering cells. There are both theoretical and practical reasons for the errors in the detection of the interfering signals. The practical reason is that the codes that we used in this simulation are comparatively simple, and thus cannot correct all the errors that an “ideal” code could eliminate. However, there is also a theoretical limit: with ideal codes, the codes in neighboring cells would be designed to have rates that achieve capacity in that cell. However, they suffer more attenuation when propagating to the neighboring cell (where they are interferers). The signal-to-noise-ratios of those signals in the neighboring cells are thus worse, so that the data rate is above the capacity of the link to the neighboring cell. Thus, correct decisions for the symbols of interfering signals might not be possible even theoretically.

Decoding of the data for interfering cells is done with the hope that this can aid in detecting the data for the desired cell. Otherwise, it is a waste of resource to do this. Moreover, wrong decision feedback can interfere with the iterative processing of the desired user, and actually worsens the performance. Thus, instead of decoding the data for all interfering cells, it makes sense to do it for just one or two strongest interfering signals and to ignore the others. The simulation results in Fig. 4.9 indicate the effectiveness of this approach. However, the performance of group IC MUD is still worse than linear MMSE.

We would expect that when we have only one dominant interfering signal, group IC MUD will outperform linear MMSE MUD. Therefore, it is worth studying the performance of group IC MUD only for the strongest interfering signal when there is one dominant interferer. The simulation results for situation (B) are shown in Fig. 4.10. We
Fig. 4.9 Performance comparison of various versions of group IC MUD when two equal-power interferers dominate

Fig. 4.10 Performance comparison of various MIMO receivers when one interferer dominates
see that group IC MUD only for the strongest interfering signal achieves the best performance, which is about 4 dB and 8 dB over Turbo-BLAST and coded V-BLAST, without MUD, respectively, and more than 2 dB over Turbo-BLAST with linear MMSE preprocessing, at 1% BLER. (Since T-BLAST (ML) assumes no advantage over T-BLAST (PIC) while having much higher complexity, we do not consider it further.)

We have noticed that group IC MUD performs the best when one interferer dominates. But when two interferers dominate that have the same power, it is no better than the simpler linear MMSE scheme. Figures 4.11 to 4.13 show that in the two-dominant-interferer scenario, when the ratio between the two largest interferer powers increases, the gap between the performance of group IC MUD and linear MMSE MUD also increases. In the view of this performance, an idea for adaptive detection arises: namely, in the case of one dominant interferer (3 dB or greater) or in the case of two dominant interferers (4 dB or greater) with the ratio between the two largest interferer power greater than 3 dB, group IC MUD could be adopted, otherwise the simple MMSE scheme could be adopted. Please note that the adaptive scheme proposed here is well suited for the corresponding setting. It should be modified when applying to other scenarios, even though the adaptive detection idea is carried on readily.

### 4.5.2 Downlink Capacity of Interference-Limited MIMO

Figures 4.14 and 4.15 give the outage capacity for interference limited MIMO when one and two interferers with equal power dominate, respectively. An upper bound (corresponding to the no-interference situation) is derived from (4.4), where the block error rate is defined as the probability that the specified spectral efficiency (8/3 bits/s/Hz
for 1/3 coded 4QAM-modulated 4 × 4 MIMO system) is not supported by the randomly generated channels. The Foschini approximation (single link capacity lower bound) is similarly derived from (4.5). For the one-dominant-interferer case, Turbo-BLAST with a parallel interference cancellation demodulation stage, with one group IC MUD (T-BLAST (PIC)+1 IC) is employed, while for the two-equal-power-dominant-interferer case, Turbo-BLAST with a parallel interference cancellation demodulation stage and with linear MMSE MUD pre-processing (T-BLAST (PIC)+MMSE) is used, as they achieve the best performance in each respective case.

![Graph](image)

**Fig. 4.11** Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 1 dB
Fig. 4.12 Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 3 dB

Fig. 4.13 Performance comparison of linear MMSE and group IC MUD when two interferers dominate with power ratio of 5 dB
The results are given for five situations: interference-free, SIR = 20, 10, 5 and 0 dB. We see that in the noise-dominating scenario (interference-free, SIR = 20 dB), the MUD capacity is excellent, even better than the Foschini approximation (Turbo-BLAST usually yields better performance than D-BLAST). Even when the SIR = 10 dB, the MUD capacity is quite close to the Foschini approximation, which is only 2-3 dB away from the exact interference-free capacity. However, when the interference gets stronger, the MUD capacity gets worse, and eventually saturates, which indicates the limitations of this method in strong interference environments and leaves ample room for possible improvement through other techniques. Note that the error floor values of Fig. 4.14 and Fig. 4.15 when SIR = 0 dB agree well with Fig. 4.10 and Fig. 4.8.

Fig. 4.14 Downlink capacity of interference-limited MIMO when one interferer dominates
Fig. 4.15 Downlink capacity of interference-limited MIMO when two interferers dominate

Fig. 4.16 Comparison of theoretical and simulated results of the capacity of interference-limited MIMO Systems with linear MMSE front end
In 5.3.3, it is proved that asymptotically (in the large dimensional system sense), the capacity of interference-limited MIMO systems with the linear MMSE preprocessing is given by

\[ C_{M-mmse} = \log \det [I + HH^H \left( \sum_i \frac{P_i}{P} H_{gi} H_{gi}^H + \frac{N}{P} I \right)^{-1}] . \]  

(4.37)

In Fig. 4.16, theoretical results of (4.37) are compared with the simulated results for the two-equal-power-dominant-interferer case (c.f. Fig. 4.15). We see that the simulated results are only 2 to 3 dB away from the capacity bound for SIR = 20–5 dB at 1% BLER, and both results exhibit the interference-limited behavior for SIR = 0 dB. The possible reasons for the gap include: 1) Our simulated system is not a large system (4 × 4 MIMO system); 2) Our Turbo-BLAST structure with the practical convolutional coding already suffers 1 to 2 dB loss in the interference-free scenario (c.f. Fig. 4.14 and Fig. 4.15). Therefore, the validity of our simulation results is verified.

### 4.5.3 Large-Scale Simulation Results

So far, the performance evaluations have been done in the framework of (4.8), where we deliberately set the SNR, SIR, and power distributions among the interferers to fixed values that represent some typical cases. In this subsection, we test the performance in the more complete model of (4.7). The receivers of interest are 1) Coded V-BLAST treating intercell interference as noise (V-BLAST), which serves as a baseline reference; 2) Turbo-BLAST with a parallel interference cancellation demodulation stage, with linear MMSE MUD (T-BLAST (PIC)+MMSE); 3) Turbo-BLAST with a parallel interference cancellation demodulation stage, with adaptive MUD detection (T-BLAST
(PIC)+ADPT); 4) Turbo-BLAST with a parallel interference cancellation demodulation stage, with the better between linear MMSE and Group IC MUD detection (T-BLAST (PIC)+IDEAL).

We again assume a 4QAM-modulated 4×4 MIMO system, with the mobile user randomly located within the cell of interest with the uniform distribution. The figure of merit is the CDF of the BLER performance for these four receivers. We collect 1000 points for this CDF profile.

### 4.5.3.1 NLOS Scenario

![CDF of BLER](image)

**Fig. 4.17 CDF of block error rate for different receivers experienced by a mobile in Rayleigh fading**

The parameters are set as in 4.5.1. The simulation results are shown in Fig. 4.17, from which we can see that 1) advanced signal processing and coding techniques substantially
improve the performance over the well-known V-BLAST technique with coding (roughly 30\% more at 1\% outage for the linear MMSE); 2) the adaptive scheme affords further gain over linear MMSE MUD (roughly 9\% more at 1\% outage for the ideal case); 3) the adaptive detection scheme illustrated in 4.5.1 approaches the ideal performance at the low BER range, which is of practical interest. Note that the threshold values of the adaptive detection scheme could be refined to get better performance in practice.

4.5.3.2 LOS Scenario

A user is randomly located as before, and the probability for LOS decreases linearly with its distance to a base station, until a "cutoff point", which is set at 300m [103]. If the signal from some base station is NLOS, the same parameters as 4.5.1 is used. Otherwise, we set the Ricean factor to

\[
K = 13 - 0.03d \text{ dB},
\]  

(4.38)

where \(d\) is the distance to some base station, and the pathloss exponent to 2. Slightly different from model (4.6), we assume no shadowing for the LOS component. Furthermore, we assume that the transmitter and receiver are positioned far apart from each other compared with the antenna spacing, so we get a rank 1 system matrix for the LOS component with energy equally distributed between real and imaginary parts [33].

The simulation results are shown in Fig. 4.18. Compared with Fig. 4.17, we see that the performance of the V-BLAST technique with coding significantly increases due to less signal fading [63]. MUD techniques with the Turbo-BLAST structure still greatly improve the system performance over the V-BLAST. But the advantage of the adaptive scheme over the linear MMSE is negligible.
This chapter has explored the downlink capacity of interference-limited MIMO cellular fading systems in fading channels. In contrast to the single-cell MIMO system considered in previous studies, where the intercell interference, when accounted for, is added to ambient Gaussian noise, we take the approach of modeling the whole downlink cellular system as a broadcast/interference channel [9], the capacity of which has long been an open question. As block fading is assumed, we are interested in outage capacity instead of the Shannon ergodic capacity. Upper bounds for this capacity are obtained from the interference-free single-link theoretical formulas. We have primarily addressed the issue of how closely we can approach those bounds without any base station
cooperation by implementation and simulation of advanced techniques. After discussing the merit of the space-time layered architecture and turbo coding/processing techniques, which come remarkably close to the ultimate capacity limits with the Gaussian ambient noise, we have considered multiuser detection for combating intercell interference. Among various multiuser detection techniques examined, linear MMSE MUD and successive interference cancellation have been shown to be feasible and effective. Successive cancellation plays a major role in network information theory from both theoretical and practical points of view. As is known, decoding of the interfering users is not always optimal except in the strong-interference case, nor is treating them as pure ambient noise optimal, except in the very-weak interference case. Based on these phenomena, we have proposed an adaptive detection idea that offers improved performance.

The success of linear MMSE processing arises, in addition to its ability to suppress interference, from its ability of producing Gaussian-like interference [80]. The observations made in [65] indicate that a receiver that uses a Gaussian-based optimal metric (which is true for our study) cannot surpass the Gaussian capacity region in the case of an ergodic additive non-Gaussian channel when Gaussian distributed codewords are selected. On the other hand, transforming the non-Gaussian interference into Gaussian-like interference guarantees the excellent performance of efficient signaling techniques well studied for AWGN channels [12], [39].

We have shown through simulation that advanced signal processing and coding techniques substantially improve the interference-limited MIMO system performance over the well-known V-BLAST techniques with coding (6-8 dB in SIR for the simplified
model, or 40% more in capacity for the cellular model, at 1% outage). We have also shown that the obtained MUD capacity is excellent in high to medium SIR environments. For the linear MMSE MUD preprocessing, our simulation results have been compared with the theoretic asymptotic capacity and their validity is verified. Our proposed techniques might be rather complex for the current systems, but will become more practically relevant in the future, as computing power at the mobile increases according to the Moore’s law. Furthermore, they are readily applicable to the base stations for uplink processing.

Finally, numerical results indicate that, due to complexity constraints and adverse environments, there is a significant performance gap between MUD capacity and interference-free capacity, especially in environments with strong interference (SIR of 5 dB or less). This indicates a need to exploit more complex schemes, such as base station cooperation (macrodiversity) for interference reduction, to enhance the system throughput.
Chapter 5

Spectral Efficiency of Multicell MIMO Systems

5.1 Introduction

Recall from Chapter 4 that a recent study by Catreux, Driessen and Greenstein [14] indicated the ineffectiveness of a MIMO system in an interference-limited environment. This seems to be related to an insufficient number of degrees of freedom (number of receive antennas) of the MIMO system to suppress the co-channel interference in contrast to the smart antenna system, where there is only one transmitted data stream in the desired cell. On the other hand, this investigation assumed a certain system structure (uncoded V-BLAST) taken from the noise-limited case, and did not try to optimize the system for interference-limited environments. Motivated by this study, we restored much of the performance of the MIMO system (to its promise in a single cell with Gaussian ambient noise) in a multi-cell structure, through application of advanced signal processing techniques. We employed a turbo space-time multiuser receiver structure for intracell communication, which essentially approaches the Shannon limit (within 1-2 dB) for an isolated cell. Furthermore, we employed another level of multiuser detection to combat the intercell interference. Among various multiuser detection techniques
examined, *group linear MMSE MUD* and *group MMSE successive cancellation* were shown to be feasible and effective. Based on these two multiuser detection schemes, each of which may outperform the other for different settings, an *adaptive multiuser detection* scheme was proposed. Simulation results indicated significant performance improvement of our approach over the well-known V-BLAST techniques with coding.

In this chapter, we study the spectral efficiency (bits/s/Hz) of the above receiver structures. We always assume optimal decoding (approximated by turbo space-time multiuser detection) for intracell communication, so the receivers are differentiated by the multiuser detection methods used to combat the intercell interference, i.e., a *single-cell detector*, the *joint optimum detector*, a *group linear MMSE detector*, a *group MMSE successive cancellation detector*, and an *adaptive multiuser detector*. We are especially interested in the asymptotic study for a large dimensional network, which is facilitated by the application of analytical results on the eigenvalue distributions of large random matrices [99], [100]. That is, we consider the limiting region where both the number of transmit antennas $K$ and receive antennas $N$ go to infinity, while their ratio remains constant. Besides its analytical convenience, the study of large system performance also has practical advantages: what revealed in the asymptotic limit is fundamental in nature, which may be concealed in the finite case by random fluctuations and other transient properties of the matrix entries; and the convergence to the asymptotic limit is rather fast with the system size.

This chapter is organized as follows. Section 5.2 presents the system model and empirical eigenvalue distributions of some large random matrices that will be useful in the sequel. In Section 5.3, formulas are derived for the optimum spectral efficiency of a
single-cell MIMO system, and spectral efficiencies of multicell MIMO systems with several optimum and sub-optimum detectors. The asymptotic study is carried out in Section 5.4. In Section 5.5, some analytical and numerical results based on these spectral efficiencies are given. Section 5.6 summarizes the chapter.

5.2 System Model

5.2.1 Single-cell and Multi-cell Communication Model

The system models we use follow those in 4.2. We simply list the mathematical models here for ease of reference. The reader is referred to Chapter 4 for details. For the single-cell model, we adopt the same mathematical model as in [105] and [42], which is given by

\[ y = Hx + n, \]

where \( y \) is the received vector, \( x \) is the transmitted signal, \( H \) is a \( N \times K \) channel matrix that captures the channel characteristics between transmit and receive antenna arrays, and \( n \) is the background noise. All through the paper, we assume \( N \geq K \), and define \( \beta = K/N \) as the system load. In the literature [129], [96], multi-cell systems are often addressed with the attractive infinite linear array model – Wyner’s model:

\[ y = Hx + \alpha H^\dagger x^\dagger + \alpha H^\dagger x^\dagger + n, \]
where only the adjacent-cell interference is taken into account, characterized by a single attenuation factor $0 \leq \alpha \leq 1$. In this chapter, we adopt the more realistic planar multi-cell model given as follows:

$$y = Hx + \sum_{i=1}^{L} \alpha_i H_{ij} x_{ji} + n,$$

(5.3)

where we assume without loss of generality $0 \leq \alpha_1 \leq \cdots \leq \alpha_L \leq 1$ with $L$ the number of effective interfering cells. In the sequel, we sometimes assign the desired cell index 0 with the understanding of $\alpha_0 = 1$ for convenience. The signal-to-noise ratio is given by $\rho = P/\sigma^2$ as before, and the signal-to-interference ratio (SIR) is given by $\mu = \frac{1}{\sum \alpha_i^2}$.

We focus mainly on the case where all the cells are identical, and all the users within each cell operate at the same rate.

### 5.2.2 Empirical Distribution of a Random Eigenvalue

Suppose $A$ is a $p \times p$ matrix with only real eigenvalues, the empirical distribution function of the eigenvalues of $A$ is defined as

$$F^A(x) = \frac{1}{p} \#(\lambda \leq x),$$

which refers to the proportion of eigenvalues of $A$ that lie below $x$. Equivalently, it can be viewed as the cumulative distribution function of a uniformly randomly selected eigenvalue of $A$. The following theorem is what we need to calculate the asymptotic spectral efficiency of MIMO systems. This theorem requires the definition of Stieltjes transform for any distribution function $G$, given as
(5.4)

for $z \in C^+ \triangleq \{ z \in C : \text{Im} z > 0 \}$.

**Theorem 5.2.1** [99]: Let $X$ be a $N \times n$ matrix containing independent and identically distributed (i.i.d.) complex entries with unit variance, and $T$ a $n \times n$ diagonal matrix, independent of $X$. Assume, almost surely, as $n \to \infty$, $F^T$ converges to a distribution function $H$, and the ratio $n/N \to c > 0$. Then, almost surely, $F^{(1/N)XX^H}$ converges to a nonrandom distribution function $G$. The Stieltjes transform $m_G(z)$ of $G$ is a unique solution to

$$m(z) = \frac{1}{-z + c \int \frac{\tau}{1 + \tau m(z)} dH(\tau)}, \quad (5.5)$$

for every $z \in C^+$.

For the special case of $\frac{1}{N}HH^H$, according to Theorem 5.2.1, the Stieltjes transform of the limiting distribution is given by

$$m_G(z) = \frac{(-1 + \beta - z) + \sqrt{-4z + (1 + z - \beta)^2}}{2z} = -\frac{1}{z} - \frac{1}{4} F\left(-\frac{1}{z}, \beta\right), \quad (5.6)$$

where

$$F(x, z) \triangleq (\sqrt{x(1 + \sqrt{z})} + 1 - \sqrt{x(1 - \sqrt{z})} + 1)^2. \quad (5.7)$$

Interestingly, the limiting distribution admits a close form expression in this case, whose probability density function is given by
\[
 f_c(x) = (1 - \beta)^+ \delta(x) + \frac{\sqrt{|x-a(\beta)|^2 |b(\beta) - x|^2}}{2\pi x},
 \]\n
where \(\delta(x)\) is a unit point mass at 0, \([x]^+ = \max\{x, 0\}\), and \(a(x) = (1 - \sqrt{x})^2\), \(b(x) = (1 + \sqrt{x})^2\). Note that (5.8) is related to the well-known fact that \(F^{(1/K)}_{\mathbf{H}^n}\) converges to the distribution function of [60], [120]

\[
 f_{\phi^+}(x) = (1 - \beta)^+ \delta(x) + \frac{\sqrt{|x-a(1/\beta)|^2 |b(1/\beta) - x|^2}}{2\pi (1/\beta)x}.
 \]\n
### 5.3 Spectral Efficiency of MIMO Systems

For the single-cell (interference-free) model (5.1) and the associated assumptions, the optimum spectral efficiency is given by [42], [105]

\[
 C_{S-opt} = \log \det \left( \frac{\sigma^2 \mathbf{I} + \frac{P}{K} \mathbf{HH}^H}{\sigma^2 \mathbf{I}} \right) = \log \det \left( \mathbf{I} + \frac{P}{K} \mathbf{HH}^H \right).
\]

Clearly, this is an upper bound for multi-cell MIMO systems. In the following, we give the spectral efficiencies of the multi-cell MIMO systems with several detectors of interest.

#### 5.3.1 Single-Cell Detector

This detector treats the intercell interference as white and Gaussian background noise, and does the “optimum” decoding for intracell communication. The observations in [65]
indicate that the spectral efficiency of this detector depends on the actual noise distribution only through its power and thus coincides with a white Gaussian noise channel with signal and noise power equal to those of the original channel. Therefore, the spectral efficiency of the single cell detector is of the same form as (5.10), with the noise spectral height replaced by \( \sigma^2 + P \sum_{i=1}^{L} \alpha_i^2 = \sigma^2 (1 + \frac{\rho}{\mu}) \). Thus, we have the following results.

**Proposition 5.3.1**: The multi-cell spectral efficiency of the desired cell MIMO system with the single-cell detector is

\[
C_{M-opt} = \log \det \left[ I + \frac{\rho}{1 + \frac{\rho}{\mu}} \frac{1}{K} HH^H \right].
\]  

(5.11)

### 5.3.2 Joint Optimum Detector

Assuming that the receiver knows the codebooks and channel information of other cells and attempts to do the joint detection, model (5.3) describes a multiple-access channel [26]. For the Gaussian multiple access channel, the capacity region is specified in [118] as (couched in the notation of the present paper)

\[
\bigcap_{I=\{1,\ldots,K(L+1)\}} \{(R_1,\ldots,R_{K(L+1)}): 0 \leq \sum_{i \in I} R_i \leq \log \det[I + \frac{\rho}{K} (H_E)_i (H_E)_i^H]\},
\]  

(5.12)

where

\[
H_E = [H, \alpha_1 H_{y_1}, \ldots, \alpha_L H_{y_L}]
\]  

(5.13)
and \((H_E)_I\), means the \(N \times |I|\) submatrix of \(H_E\) obtained by striking out the columns whose indices do not belong to \(I\), with \(|I|\) the cardinality of \(I\). Here \(R_i \sim R_K\) denotes the data rates of the MIMO system in the desired cell, \(R_{K+1} \sim R_{2K}\) refers to that of the first interfering cell with attenuation factor \(\alpha_i\), and so on. The following proposition is a specific application of (5.12).

**Proposition 5.3.2**: The partial sum spectral efficiency of the multi-cell MIMO systems with the joint optimum detector is

\[
S R_{M-opt}^{(j)} = \log \det \left[ I + \frac{P}{K} (H_E)_J (H_E)_J^H \right],
\]

where \(J \subset \{0,1,\ldots,L\}\) denotes the set of cells of interest (cell 0 is the desired cell), and \((H_E)_J\) means the \(N \times |J| \times K\) submatrix of \(H_E\) obtained by striking out the channel matrices of the interfering cells whose indices do not belong to \(J\).

**Corollary 5.3.1**: The multi-cell spectral efficiency of the desired cell MIMO system with the joint optimum detector is the same as that of the single-cell MIMO systems,

\[
C_{M-opt} = S R_{M-opt}^{(0)} = C_{S-opt}.
\]

**Comments**: Note that by employing a joint maximum likelihood detector at the receiver of the desired cell, even though the data from the interfering cells may not be detected correctly (the data rate of interfering cells exceeds the link capacity viewed at the receiver of the desired cell), the data of the desired cell will be detected with high fidelity, as long as the data rate of the desired cell is below its single-cell capacity.
5.3.3 Group Linear MMSE Detector

As we mentioned in Chapter 4, joint maximum likelihood detection for multi-cell MIMO is impractical for most current applications due to its complexity. With no intention to detect the data from the interfering cells, group linear MMSE MUD is one of the most favorable techniques to suppress the intercell interference. The detection process is to first apply the weight matrix

\[
W = \left( HH^H + \sum_i \alpha_i^2 H_{ji} H_{ji}^H + \frac{K}{\rho} I \right)^{-1} H
\]

(5.16)
to the received signal (5.3) to combat the intercell interference, and then to optimally detect the data of the desired cell.

When there is no attempt to detect the data from the interfering cells, model (5.3) should be understood as a non-Gaussian interference channel in general. Even though the non-Gaussian channel capacity is higher than that of the Gaussian channel with the same input power constraint, it can hardly be approached due to the following reasons. First, the noise characteristics are either not known exactly (e.g., in the interference channel) or are too expensive to exploit. Second, capacity-approaching signaling for non-Gaussian channels is far less well understood than that for the Gaussian channel. Therefore, the Gaussian-based optimal metric is widely used in receiver design. It is shown in [65] and [64] that, in the case of an ergodic additive non-Gaussian channel, when Gaussian codebooks are used, a receiver that exploits a Gaussian-based optimal metric can not surpass the Gaussian capacity, defined as follows.

Definition 5.3.1: The Gaussian capacity of an additive non-Gaussian channel is the capacity of an additive Gaussian channel with the same noise covariance matrix.
According to this definition, the Gaussian capacity of (5.3) is given by

\[ C_{M \rightarrow G} = \log \det \left[ I + \frac{P}{K} HH^H \Sigma^{-1} \right], \quad (5.17) \]

where

\[ \Sigma = \sum_{i=1}^{L} \alpha_i^2 \frac{P}{K} H_{ii}^H H_{ii} H^H + \sigma^2 I. \quad (5.18) \]

Even though, theoretically, this mismatched decoding results in a capacity equal to the Gaussian capacity of the original non-Gaussian interference channel [65], in practice, as the efficient capacity-approaching turbo-style probabilistic decoding is increasingly used, where the probabilistic information is produced with a Gaussian assumption, this mismatched decoding usually results in a significant capacity loss. On the other hand, as we have noted, the distribution of the residual interference plus noise at the output of a linear MMSE multiuser detector is well approximated by a Gaussian distribution [80], [133], which guarantees the excellent performance of efficient signaling and decoding techniques well studied for AWGN channels [12], [39]. The following proposition verifies the effectiveness of linear MMSE preprocessing.

**Proposition 5.3.3**: The group linear MMSE detector asymptotically achieves the Gaussian capacity

\[ C_{M \rightarrow \text{mmse}} = C_{M \rightarrow G} = \log \det \left[ I + \frac{P}{K} HH^H \Sigma^{-1} \right]. \quad (5.19) \]

**Proof**: After linear MMSE filtering with (5.16) the system model can be represented as

\[ y' = W^H H \cdot x + \eta, \]
where \( \eta \) is Gaussian distributed with covariance matrix of \( \mathbf{W}^H \mathbf{\Sigma} \mathbf{W} \). This is verified in [133] as \( N \to \infty \). The spectral efficiency of this model is given by

\[
C_{M-MMSE} = \log \det[\mathbf{I} + \frac{P}{K} \mathbf{W}^H \mathbf{H} \mathbf{H}^H \mathbf{W}(\mathbf{W}^H \mathbf{\Sigma} \mathbf{W})^{-1}].
\] (5.20)

With (5.16) and (5.18), it is easy to verify that

\[
\mathbf{W}^H \mathbf{\Sigma} \mathbf{W} = \frac{P}{K} (\mathbf{I} - \mathbf{W}^H \mathbf{H}) \mathbf{W}^H \mathbf{H}.
\]

Assume \( K \leq N \) throughout, and define \( \mathbf{Q} = \frac{P}{K} \mathbf{H}^H \mathbf{\Sigma}^{-1} \mathbf{H} \), it is shown that the probability that \( \mathbf{Q} \) is non-singular goes to 1 as \( N \to \infty \) [120]. Then

\[
\mathbf{W}^H \mathbf{H} = \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \frac{K}{P} \mathbf{\Sigma}^{-1}) \mathbf{H}
\]

\[
= \mathbf{H}^H \left( \frac{P}{K} \mathbf{\Sigma}^{-1} - \left( \frac{P}{K} \mathbf{\Sigma}^{-1} \right) \mathbf{H} \mathbf{\Delta}^{-1} \mathbf{H}^H \left( \frac{P}{K} \mathbf{\Sigma}^{-1} \right) \right) \mathbf{H},
\]

with \( \Delta = \mathbf{I} + \mathbf{H}^H \left( \frac{P}{K} \mathbf{\Sigma}^{-1} \right) \mathbf{H} = \mathbf{I} + \mathbf{Q} \), by the matrix inverse formula. It then follows that \( \mathbf{W}^H \mathbf{H} = \mathbf{Q} - \mathbf{Q} (\mathbf{I} + \mathbf{Q})^{-1} \mathbf{Q} = (\mathbf{I} + \mathbf{Q})^{-1} \mathbf{Q} \), and \( \mathbf{I} - \mathbf{W}^H \mathbf{H} = \mathbf{I} - (\mathbf{I} + \mathbf{Q})^{-1} \mathbf{Q} = (\mathbf{I} + \mathbf{Q})^{-1} \), both of which are invertible asymptotically.

Therefore (note \( \mathbf{W}^H \mathbf{H} = \mathbf{H}^H \mathbf{W} \))

\[
C_{M-MMSE} = \log \det[\mathbf{I} + \mathbf{W}^H \mathbf{H} (\mathbf{I} - \mathbf{W}^H \mathbf{H})^{-1}]
\]

\[
= \log \det[(\mathbf{I} - \mathbf{W}^H \mathbf{H})^{-1}]
\]

\[
= \log \det[(\mathbf{I} - (\mathbf{I} + \mathbf{Q})^{-1} \mathbf{Q})^{-1}]
\]

\[
= \log \det[\mathbf{I} + \mathbf{Q}] = C_{M-MMSE}.
\] (5.21)

QED.
By Proposition 5.3.3, we have the following consequence.

Corollary 5.3.2: The multi-cell spectral efficiency of the desired MIMO system with the linear MMSE detector is

\[ C_{M \rightarrow m \text{mse}} = SR_{M \rightarrow opt}^{(0,-\cdots,-\cdots)} - SR_{M \rightarrow opt}^{(0,-\cdots,-\cdots)}. \]  

(5.22)

Proof: By (5.18),

\[ \Sigma = \sum_{i=1}^{L} \alpha_i^2 \frac{P}{K} H_{fi} H_{fi}^H + \sigma^2 I \]

\[ = \frac{P}{K} (H_E H_E^H - HH^H) + \sigma^2 I. \]

So

\[ \frac{P}{K} (H_E H_E^H - HH^H) = \frac{1}{\sigma^2} \Sigma - I \]

and

\[ \frac{P}{K} H_E H_E^H = \frac{P}{K} HH^H + \frac{1}{\sigma^2} \Sigma - I \]

Therefore,

\[ I + \frac{P}{K} HH^H \Sigma^{-1} = I + \frac{P}{K} H_E H_E^H \]

\[ \frac{1}{I + \frac{P}{K} (H_E H_E^H - HH^H)}. \]  

(5.23)

Compare (5.14) and (5.19), (5.22) follows.

QED.
In general, if we partition the cells in two groups, applying the linear MMSE decoding to one of them while treating the other as interfering cells, the sum spectral efficiency is exactly analogous to (5.22). Thus, we have the following.

*Corollary 5.3.3:* The partial sum spectral efficiency of multi-cell MIMO systems with the generalized group linear MMSE detector is

\[
SR_{M-mmse}^{(J)} = SR_{M-opt}^{(0,1,...,L)} - SR_{M-opt}^{(\bar{J})},
\]

where \( J \subset \{0,1,...,L\} \) denotes the set of cells of interest (cell 0 is the desired cell), and \( \bar{J} \) is the complement of \( J \) in \( \{0,1,...,L\} \).

*Comments:* Note that \( SR_{M-mmse}^{(0)} = C_{M-mmse} \), while \( SR_{M-mmse}^{(0,1,...,L)} = SR_{M-opt}^{(0,1,...,L)} \). Within the set of cells of interest \( J \), while treating the interference from other cells as Gaussian background noise (due to MMSE processing), we can similarly define a multiple access capacity region as Proposition 5.3.2. Denote a set \( K \subset J \), then the partial sum spectral efficiency of this set is given by

\[
SR_{M-mmse}^{(J,K)} = \log \det \left[ I + \frac{P}{K} (H_K) K (H_K) K ^H \Sigma_J ^{-1} \right],
\]

where \( \Sigma_J = \sum_{i,j} \alpha_i \alpha_j \frac{P}{K} H_{ji} H_{ji} ^H + \sigma^2 I \). Similarly (5.25) can be represented by

\[
SR_{M-mmse}^{(J,K)} = SR_{M-opt}^{(K \cup \bar{J})} - SR_{M-opt}^{(\bar{J})},
\]

where \( J \subset \{0,1,...,L\} \) denotes the set of cells of interest (cell 0 is the desired cell), and \( \bar{J} \) is the complement of \( J \) in \( \{0,1,...,L\} \).
5.3.4 Group MMSE Successive Cancellation Detector

As noted previously, since joint maximum likelihood detection for multi-cell MIMO is highly complex, while linear MMSE MUD is limited in its interference cancellation capability, especially when $\beta \geq 1$ [120], non-linear multiuser detection often provides a favorable tradeoff between performance and complexity. Group MMSE successive cancellation is one such technique, in which information bits are detected by group linear MMSE detection cell by cell, with the interference from previously detected cells already being subtracted. Although successive cancellation does not result in maximum-likelihood decisions, it becomes asymptotically optimal as the error probability of intermediate decisions vanishes with code block length [116]. With the assumption of perfect cancellation, the following lemma shows the optimality of group MMSE successive interference cancellation.

**Lemma 5.3.1:** Multi-cell MIMO systems with the group MMSE successive cancellation detector applied to the set of cells of interest $J \subset \{0,1,\ldots,L\}$ achieves some vertices of the capacity region given by (5.25) and (5.26).

**Proof:** Suppose $J = \{i_1,\ldots,i_J\}$. Without loss of generality, suppose also that the cells are decoded in that order. Following Corollary 5.3.2, and assuming perfect cancellation of the interference from the detected cells, we have

$$C_{M-sic}^{\{i_1\}} = SR_{M-opt}^{\{0,1,\ldots,i_1\}} - SR_{M-opt}^{\{0,1,\ldots,i_1\}/i_1}$$

$$C_{M-sic}^{\{i_2\}} = SR_{M-opt}^{\{0,1,\ldots,i_1\}/i_1} - SR_{M-opt}^{\{0,1,\ldots,i_1\}/(i_1,i_2)}$$

$$\vdots$$
\begin{align}
C_{M-\text{sic}}^{(i_j)} &= SR_{M-\text{opt}}^{(0,1,\ldots,J-1)} - SR_{M-\text{opt}}^{(0,1,\ldots,J-1) - (i_j)}. 
\tag{5.27}
\end{align}

Equivalently, the above $|J|$ equalities can be reformulated as

\begin{align}
SR_{M-\text{sic}}^{(J)} &= C_{M-\text{sic}}^{(i_j)} = SR_{M-\text{opt}}^{(i_j)} - SR_{M-\text{opt}}^{(J)} \\
SR_{M-\text{sic}}^{(J,\ldots,i_{j_1})} &= C_{M-\text{sic}}^{(i_{j_1})} + C_{M-\text{sic}}^{(i_{j_2})} = SR_{M-\text{opt}}^{(i_{j_1})} - SR_{M-\text{opt}}^{(J)} \\
&\vdots \\
SR_{M-\text{sic}}^{(J)} &= C_{M-\text{sic}}^{(i_{j_1})} + \cdots + C_{M-\text{sic}}^{(i_{j_k})} = SR_{M-\text{opt}}^{(i_{j_1})} - SR_{M-\text{opt}}^{(J)}. \tag{5.28}
\end{align}

On comparing (5.28) with (5.26), we see that one of the vertices of the Gaussian multiple access capacity region is achieved.

As the MMSE successive interference cancellation detector has $|J|$ different orders of detection, $|J|$ vertices of the capacity region in the $R^{[J]}$ domain can be achieved. As the capacity region of the multiple-access channel is convex, by timesharing, the MMSE successive interference cancellation detector can thus achieve the capacity region spanned by these vertices.

\textbf{QED.}

\textbf{Corollary 5.3.4:}

\begin{align}
SR_{M-\text{sic}}^{(J)} &= \sum_{i \in J} C_{M-\text{sic}}^{(i)} = SR_{M-\text{opt}}^{(0,1,\ldots,J-1)} = SR_{M-\text{mmse}}^{(J)} - SR_{M-\text{mmse}}^{(0,1,\ldots,J-1)}. 
\tag{5.29}
\end{align}

\begin{align}
SR_{M-\text{sic}}^{(0)} &= SR_{M-\text{mmse}}^{(0)} = C_{M-\text{mmse}}. 
\tag{5.30}
\end{align}

\begin{align}
SR_{M-\text{sic}}^{(0,1,\ldots,L_1)} &= SR_{M-\text{mmse}}^{(0,1,\ldots,L_1)} = SR_{M-\text{opt}}^{(0,1,\ldots,L_1)}. 
\tag{5.31}
\end{align}
To achieve some vertex of the capacity region of the Gaussian multiple access channel that corresponds to the maximum rate for the desired cell (single cell capacity), Lemma 5.3.1 suggests a detection order that puts the detection of the desired cell last. For example, suppose the cells are detected in the order of \( \{L, L-1, \ldots, 0\} \), the achieved capacity vertex is given by

\[
(R_{M-sic}^{(L)}, \ldots, R_{M-sic}^{(1)}, R_{M-sic}^{(0)}),
\]

with

\[
R_{M-sic}^{(l)} = SR_{M-opt}^{(0)} - SR_{M-opt}^{(l-1)}, \quad l = 0, 1, \ldots, L.
\]

Clearly,

\[
R_{M-sic}^{(0)} = SR_{M-opt}^{(0)} = C_{S-opt}.
\]

Note that in practice, the success of interference cancellation relies highly on the correct detection of interference. In adverse environments where we cannot get good estimates of interference, successive interference cancellation schemes will worsen the performance instead of improving it. Therefore, to achieve the optimum capacity of (5.34), the MMSE successive interference cancellation detector implicitly requires that the data rates of other interfering cells satisfy (5.33). This is impractical, as it requires not only joint signaling but also puts the desired cell in a superior position. What is of practical interest is that all cells are autonomous with identical data rates. This identical rate tuple can in general be achieved by timesharing. With this restriction, the following proposition shows the limitation of the MMSE successive interference cancellation detector.
Proposition 5.3.4: Assuming that the same data rate is employed in each cell, and that $0 \leq \alpha_L \leq \cdots \leq \alpha_1 \leq 1$, the multi-cell spectral efficiency of the desired cell MIMO system with the group MMSE successive cancellation detector applied to the cells $J = \{0, 1, \ldots, l\}$ is

$$C_{M \text{-- sic}}^{(l)} = \min_{k=1, \ldots, j+1} \left( \frac{SR_{M \text{-- opt}}^{(j+1, \ldots, l)} - SR_{M \text{-- opt}}^{(j, \ldots, l)}}{k} \right), l = 0, 1, \ldots, L. \quad (5.35)$$

Proof: As the same data rate is employed in each cell, we have

$$SR_{M \text{-- sic}}^{(j, k)} = |K| C_{M \text{-- sic}}^{(l)}. \quad (5.36)$$

We know from Lemma 5.3.1 that the achievable rate-tuples are constrained by (5.26). Due to the condition $0 \leq \alpha_L \leq \cdots \leq \alpha_1 \leq 1$, we can rule out most of the constraints. For single rate, we have

$$C_{M \text{-- sic}}^{(l)} \leq \min_{j \in J} \left( SR_{M \text{-- opt}}^{(j)} - SR_{M \text{-- opt}}^{(j+1, \ldots, l)} \right) \leq SR_{M \text{-- opt}}^{(j, \ldots, l)} - SR_{M \text{-- opt}}^{(j+1, \ldots, l)}. \quad (5.37)$$

Similarly, for sum rates of two cells, we have

$$2C_{M \text{-- sic}}^{(l)} \leq \min_{j, k \in J} \left( SR_{M \text{-- opt}}^{(j, k)} - SR_{M \text{-- opt}}^{(j+1, \ldots, l)} \right) \leq SR_{M \text{-- opt}}^{(j, \ldots, l)} - SR_{M \text{-- opt}}^{(j+1, \ldots, l)}. \quad (5.38)$$

So,

$$C_{M \text{-- sic}}^{(l)} \leq \frac{SR_{M \text{-- opt}}^{(j+1, \ldots, l)} - SR_{M \text{-- opt}}^{(j, \ldots, l)}}{2}. \quad (5.39)$$

By a similar argument, (5.35) follows readily.
Comments: Note that to jointly detect \( l + 1 \) cells, including cell 0, the set of 
\[ J = \{0, 1, \ldots, l\} \] is optimal, as any other choice can only lower the capacity (5.35), due to
\[ 0 \leq \alpha_l \leq \cdots \leq \alpha_1 \leq 1. \] Also note that \( C_{M\text{-sic}}^{(0)} = C_{M\text{-mmse}} \).

5.3.5 Adaptive Multiuser Detector

From Proposition 5.3.4, we see that the group MMSE successive cancellation detector is 
not necessarily better than the simpler group linear MMSE detector. Similarly, it is not 
always better to try to detect more cells. These observations are confirmed in Chapter 4. 
An (ideal) adaptive detector will always assume the best performance among linear 
MMSE MUD and various partial or full interference cancellation detectors. This detector 
can be approximated by a receiver that chooses different detection schemes according to 
some thresholds determined by experiments (see 4.5.1).

Proposition 5.3.5: The multi-cell spectral efficiency of the desired cell MIMO system 
with the adaptive multiuser detector is
\[
C_{M\text{-adpt}} = \max_{l = 0, \ldots, L} C_{M\text{-sic}}^{(l)}. \tag{5.40}
\]

5.4 Asymptotic Study

From the previous section, we see that all formulas of interest are expressed in the form 
of (5.14). In this section, we obtain a nonrandom expression for (5.14) in the limiting 
region. To this end, we rewrite (5.14) as
\[
\lim_{N \to \infty} SR^{(j)}_{H_{\text{opt}}} = \lim_{N \to \infty} \log \det \left[ I + \frac{\rho}{K} (H_E)_{j,j} (H_E)_{j,j}^H \right] \\
= \lim_{N \to \infty} \log \det \left[ I + \frac{1}{\beta N} (H_E)_{j,j} (H_E)_{j,j}^H \right] \\
= N \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \log \left( 1 + \frac{\rho}{\beta} \lambda_i \right) \\
= NE \left\{ \log \left( 1 + \frac{\rho}{\beta} \lambda \right) \right\}
\]

(5.41)

where \( \{\lambda_i\} \) are the eigenvalues of

\[
\frac{1}{N} (H_E)_{j,j} (H_E)_{j,j}^H = \frac{1}{N} [H_{t_i}, \ldots, H_{t_j}] \begin{bmatrix}
\alpha_{t_i}^2 \\
\vdots \\
\alpha_{t_j}^2
\end{bmatrix} [H_{t_i}, \ldots, H_{t_j}]^H.
\]

(5.42)

To calculate (5.41), notice that

\[
C(\gamma) = E\{\log(1 + \gamma \lambda)\}
\]

(5.43)

is an increasing function of \( \gamma \) with \( C(0) = 0 \), so we can express (5.43) as

\[
C(\gamma) = \int_0^\gamma \frac{d}{dx} C(x) dx \\
= \int_0^\gamma E \left\{ \frac{\lambda}{1 + x \lambda} \right\} dx.
\]

(5.44)

Further notice that (5.42) conforms with the conditions of Theorem 5.2.1, so the empirical eigenvalue distribution of \( \frac{1}{N} (H_E)_{j,j} (H_E)_{j,j}^H \) converges in distribution to a nonrandom distribution function \( Q \), whose Stieltjes transform \( m_Q(z) \) is a unique solution to

...
\[ m_Q(z) = \frac{1}{-z + |J|/\beta \int \frac{\tau}{1 + \tau m_Q(z)} \, dH(\tau)} \]
\[ \quad = \frac{1}{-z + \beta \sum_{j=1}^{|J|} \frac{\alpha_j^2}{1 + \alpha_j^m(z)}} \]  

(5.45)

By (5.4), we have
\[ m_Q(z) = E \left\{ \frac{1}{\lambda - z} \right\}. \]  

(5.46)

So,
\[ E \left\{ \frac{\lambda}{1 + x\lambda} \right\} = \frac{x - m_Q\left(\frac{1}{x}\right)}{x^2}. \]  

(5.47)

Therefore,
\[ \lim_{N \to \infty} SR^{(s)}_{m,\text{opt}} = N \int_0^{\rho/\beta} \frac{x - m_Q\left(\frac{1}{x}\right)}{x^2} \, dx, \]  

(5.48)

where \( m_Q(x) \) is an implicit solution of (5.45).

The optimum spectral efficiency of single-cell MIMO systems (5.10) can be calculated in the same way as above. However, due to the expression of (5.9), a closed-form expression can be derived as follows. In the limiting region, we can express (5.10) as

\[ \lim_{N \to \infty} C_{S,\text{opt}} = N \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \log(1 + \rho\lambda_i) \]
\[ = NE \{ \log(1 + \rho\lambda) \}. \]  

(5.49)
where \( \{\lambda_i\} \) are the eigenvalues of \( \frac{1}{K} \HH^H \), whose limiting probability density function is given by (5.9). Using the same differentiation-integration strategy, we have

\[
\lim_{N \to \infty} C_{S-opt} = N\int_0^\rho E \left\{ \frac{\lambda}{1 + x\lambda} \right\} dx
= N\int_0^\rho \left( 1 - E \left\{ \frac{1}{1 + x\lambda} \right\} \right) dx. \tag{5.50}
\]

Fortunately, \( E \left\{ \frac{1}{1 + x\lambda} \right\} \) can be calculated explicitly:

\[
E \left\{ \frac{1}{1 + x\lambda} \right\} = \int \frac{1}{1 + x\lambda} f_{1/\beta}(\lambda)d\lambda
= (1 - \beta) + \frac{1}{2\pi(1/\beta)} \int_{a(1/\beta)}^{b(1/\beta)} \frac{1}{1 + x\lambda} \sqrt{(\lambda - a(1/\beta))[b(1/\beta) - \lambda]} d\lambda
= (1 - \beta) + \beta \int_{a(\beta)}^{b(\beta)} \frac{1}{1 + \frac{x}{\beta}\omega} \sqrt{(\omega - a(\beta))[b(\beta) - \omega]} \cdot \frac{d\omega}{2\pi\beta}\omega \tag{5.51}
= (1 - \beta) + \beta \left( 1 - \frac{1}{4x} F \left( \frac{x}{\beta}, \beta \right) \right)
= 1 - \frac{\beta}{4x} F \left( \frac{x}{\beta}, \beta \right),
\]

where the second-to-last definite integral is derived in [49], [120], and the function \( F(x, z) \) is given in (5.7). Finally,

\[
\lim_{N \to \infty} C_{S-opt} = N\int_0^\rho \frac{\beta}{4x^2} F \left( \frac{x}{\beta}, \beta \right) dx
= N\int_0^\rho \frac{1}{4y^2} F (y, \beta) dy \tag{5.52}
= NG \left( \frac{\rho}{\beta}, \beta \right),
\]
where

$$G(x, z) = z \log(1 + x - \frac{1}{4} F(x, z)) + \log(1 + xz - \frac{1}{4} F(x, z)) - \frac{\log e}{4x} F(x, z).$$

(5.53)

Similarly, we have

$$\lim_{N \to \infty} C_{M-\text{opt}} = NG\left(\frac{\rho}{\rho / \mu} \beta; \beta\right).$$

(5.54)

### 5.5 Some Analytical and Numerical Results

In this section, some analytical and numerical results are given as applications of the above derived formulas, from which we can get some insights into the behavior of multicell MIMO systems. Unless otherwise specified, we assume such a multicell model having four interferers in two groups of two, in which one group is 6 dB stronger than the other while the users within each group have the same power. This roughly reflects the essential reality as interference from the two farthest adjacent cells can typically be ignored, and simulation results in Chapter 4 verify that the power of the two strongest users usually dominates. Therefore, we assume the following parameters for (5.3): $L = 4$ and

$$
\begin{align*}
\alpha_1^2 &= \frac{1}{\mu} \frac{\gamma}{1 + \gamma} \frac{\beta_1}{1 + \beta_1}, \\
\alpha_2^2 &= \frac{1}{\mu} \frac{\gamma}{1 + \gamma} \frac{1}{1 + \beta_1}, \\
\alpha_3^2 &= \frac{1}{\mu} \frac{1}{1 + \gamma} \frac{\beta_2}{1 + \beta_2}, \\
\alpha_4^2 &= \frac{1}{\mu} \frac{1}{1 + \gamma} \frac{1}{1 + \beta_2}.
\end{align*}
$$

(5.55)
with $\gamma = 4, \beta_1 = 1, and \beta_2 = 1$. Recall that $\mu$ is the SIR. Further, as spectral efficiencies of MIMO systems grow linearly with the number of receive antennas $N$, in our study we are mainly interested in normalized spectral efficiencies per receive antenna. The reader should keep this in mind when interpreting the overall MIMO capacity from the following figures.

### 5.5.1 Approximate Formula

The exact formula of (5.48) requires numerical fixed-point solutions of (5.45) and the definite integral of (5.48), which is fairly complex. So, an approximating formula of (5.14) is explored here. The key idea is to approximate $$(H_E)_J(H_E)_J^H$$ as

$$
(H_E)_J(H_E)_J^H \approx \sum_{i=1}^{\beta} |J| \mathbf{H}'(\mathbf{H}')^H,
$$

(5.56)

where $\mathbf{H}'$ denotes an $N \times K |J|$ random matrix with i.i.d. normalized complex Gaussian entries. In this way, a closed-form formula similar to (5.52) can be obtained. Note that even though we assume $\beta \leq 1$ all through the chapter, we should discern here whether $\beta \leq \frac{1}{|J|}$, which determines whether the empirical eigenvalue distribution of $\frac{1}{K \sum_{i=J} \alpha_i^2} (H_E)_J(H_E)_J^H$ has a mass point at 0 (see (5.9)). Following the same lines as (5.52), we get the following approximate formula for (5.14), which coincides for all values of $\beta$: 

\[ \lim_{N \to \infty} SR_{M_{\text{opt}}}^{(i)} \approx N \beta |J| G \left( \rho \sum_{i \in J} \alpha_i^2, 1/\beta |J| \right) \]

\[ = NG \left( \rho \sum_{i \in J} \alpha_i^2 |J|, \beta |J| \right). \]  

(5.57)

We will see in the following that (5.57) gives a good enough approximation for a wide range of parameter settings. It tends to overestimate when \( \beta \) is small or there is great discrepancy within the set of \( \{\alpha_i\} \) of interest. Even in this case, (5.57) roughly exhibits the same behavior as (5.48), and thus is still useful for theoretical analysis.

### 5.5.2 Interference-Limited Behavior

Clearly, the single-cell detector is interference limited, which can be verified by (5.54):

\[ \lim_{\rho \to \infty} G(\rho/(1 + \rho/\mu) \beta, \beta) = G(\mu/\beta, \beta). \]  

(5.58)

Even though for \( \beta = 1 \), the group linear MMSE detector is also interference limited (see Chapter 4), we noted that it is due to the lack of sufficient degrees of freedom at the receiver to suppress the co-channel interference. We believe that if \( \beta \) is sufficiently small, the group linear MMSE detector is not interference limited. This is verified in Figs 5.1 (a)-(c). Here the SIR is set to be 0 dB, indicating a strong interference environment. It is observed in these figures that the spectral efficiency of the single cell upper bound (see (5.52)) and that of the single-cell detector (see (5.54)) decrease as the system load \( \beta \) decreases. The single-cell detector is interference-limited, with the limiting value being given by (5.58). The spectral efficiency of the group linear MMSE detector, both the exact (see (5.24) and (5.48)) and the approximate (see (5.24) and (5.57)), however,
increases as $\beta$ decreases, when the SNR is sufficiently large. Furthermore, when the system load decreases to $1/5$, the group linear MMSE detector is not interference limited. Comparing Figs 5.1 (a), (c) with (d), where a more favorable SIR = 5 dB is experienced, we can see that all the multicell spectral efficiencies increase as SIR increases. However, due to the interference-limited nature, the spectral efficiency of the group linear MMSE detector with system load 1 is outperformed by the same detector with system load of $1/5$ at a much worse SIR, when the SNR is sufficiently large. This observation is helpful for multicell MIMO system design. Finally, we observe that the approximate formula well matches the exact one, thus providing a valuable tool for analysis.

Let us turn to the normalized approximate formula, given as

$$\lim_{N \to \infty} C_{M-\text{mmse}} / N \approx R(\rho, \mu, \beta) \triangleq G\left(\frac{\rho(1+1/\mu)}{\beta(L+1)}, \beta(L+1)\right) - G\left(\frac{\rho(1/\mu)}{\beta L}, \beta L\right), \quad (5.59)$$

to study the interference-limited behavior of the group linear MMSE detector. In the Appendix I, we show that

$$\lim_{\rho \to \infty} R(\rho, \mu, \beta) = \log(1+\mu) + (1-\beta(L+1)) \log\left(1 - \frac{1}{\beta(L+1)}\right) + (\beta L - 1) \log\left(1 - \frac{1}{\beta L}\right) \quad \text{when} \quad \beta > \frac{1}{L}, \quad (5.60)$$

and

$$\lim_{\rho \to \infty} R(\rho, \mu, \beta) = \infty \quad \text{when} \quad \beta \leq \frac{1}{L+1}. \quad (5.61)$$
The analytical results of (5.60) and (5.61) agree with the numerical results of Fig. 5.1 very well. The “magic” number 1/5 is not found by luck but rather is determined by the system behavior.
CHAPTER 5. SPECTRAL EFFICIENCY OF MULTICELL MIMO SYSTEMS

(b) SIR = 0 dB, System Load = 1/2

(c) SIR = 0 dB, System Load = 1/5
Fig. 5.1 Study of interference-limited behavior of various multicell MIMO spectral efficiencies

5.5.3 Adaptive Detection

We continue to study the behavior of the group MMSE successive cancellation detector (see (5.35)) and the adaptive multiuser detector (see (5.40)). The asymptotic spectral efficiencies given in this subsection are calculated with the exact formula (5.48). In the following figures, we use “Group MMSE SC-\( l \)” to denote the multi-cell spectral efficiency of the desired MIMO system with the group MMSE successive cancellation detector applied to the cells \( J = \{0,1,\ldots,l\} \). The group linear MMSE detector corresponds to “Group MMSE SC-1”, and the adaptive detector achieves the best among these detectors.
From Figs 5.2 (a) and (b) we see that, for a fairly high SIR, the simpler group linear MMSE detector is the best; but in a strong interference environment, group MMSE successive cancellation proves to be useful. This verifies the well-known fact that, detection of the interfering users is only optimal in the strong-interference case; for weak interference, it is better to simply treat them as ambient noise. It is also observed that, in the sufficiently high SNR scenario, the group MMSE successive cancellation detector applied to all the cells eventually stands out; but for other cases, trying to detect more cells actually lowers the possible achieved capacity. This is more evident in Fig. 5.2 (c), where a one-dominant-interferer scenario with $\gamma = 3.5, \beta_1 = 6, \beta_2 = 1$ (the power of the strongest interferer is 3 dB higher than the power sum of the rest interferers) is assumed. We find that for low to medium SNR, detection of only the strongest interferer is the best, while full MMSE successive cancellation is the best in the high SNR regime.

In Fig. 5.3 (a), we show the spectral efficiency of the ideal adaptive detector for different SNR and SIR scenarios. We assume the model of (5.55) with $\gamma = 4, \beta_1 = 1, \beta_2 = 1$, and with system load 1. We also show the single cell upper bound and that of the single-cell detector for reference. We see that multiuser detection across
CHAPTER 5. SPECTRAL EFFICIENCY OF MULTICELL MIMO SYSTEMS 174

(a) SIR = 5 dB, System Load = 1

(b) SIR = 0 dB, System Load = 1
Fig. 5.2 Spectral efficiency comparison of group linear MMSE detector, group MMSE successive cancellation detector, and adaptive detector
the cell is most useful in the strong interference environment. Its advantage over the single-cell detector diminishes as SIR increases. However, note that in practice, while the group linear MMSE capacity (which is also the adaptive detector capacity in high SIR, see Fig. 5.2 (a)) can be closely approached, there is no effective way to approach the single-cell detector capacity due to the non-Gaussian nature of the channel. Also note that a small difference in the spectral efficiency per receive antenna may become substantial in large systems. In the strong interference environment, even though multiuser detection across the cell brings substantial gain over single-cell detection, there is a substantial gap between the achievable capacity and the single cell upper bound. This is not the case for sufficiently low system load, as is shown in Fig. 5.3 (b). However, as the system load is reduced, the achievable capacity is also reduced. In the strong interference environment, it is possible to improve the system performance by reducing the system load and
increasing the signal power, though this may be costly. In the low interference environment, however, it is better to have a high system load and a group linear MMSE receiver.

5.6 Summary

In this chapter, the spectral efficiency of multicell MIMO systems with several MUD detectors has been studied, among which are a single-cell detector, the joint optimum detector, a group linear MMSE detector, a group MMSE successive cancellation detector, and an adaptive multiuser detector. The large-system asymptotic (non-random) expressions for these spectral efficiencies have also been explored.

As applications of these theoretical bounds, we have verified the following observations from Chapter 4 about multicell MIMO systems. Group linear MMSE detection and group MMSE successive cancellation are two effective techniques to combat the co-channel interference, each of which may outperform the other for different settings. Based on this observation, an adaptive detector was developed to attempt to achieve the better of the two. For full system load, multiuser detection across the cell is most useful in strong interference environments, offering substantial gain over traditional single-cell detection. However, there is still a remarkable gap between the achievable capacity and the single cell upper bound in this case.

Further, conditions for non-interference-limited behavior of the group linear MMSE detector have been found. Based on this result, it is suggested that with sufficiently low
system load $\beta \leq \frac{1}{L+1}$, where $L$ is the number of effective interfering cells, better performance than that of the fully loaded system may be attained in the strong interference environment with sufficiently large signal power. If the interference is weak, then a high-load system with fairly simple group linear MMSE detection has the advantage.

**Appendix I**

First let us examine the limiting behavior of the three parts of $G(x, z)$ in (5.53):

$$\lim_{x \to \infty} \frac{F(x, z)}{4x} = \min(z, 1), \quad (5.62)$$

$$\lim_{x \to \infty} 1 + x - \frac{1}{4} F(x, z) = \lim_{x \to \infty} \frac{1}{1 - F(x, z)/4x} = \frac{1}{1 - \min(z, 1)/z}, \quad (5.63)$$

$$\lim_{x \to \infty} 1 + xz - \frac{1}{4} F(x, z) = \lim_{x \to \infty} \frac{1}{1 - F(x, z)/4x} = \frac{1}{1 - \min(z, 1)}. \quad (5.64)$$

We see that when $z \leq 1$, (5.63) goes to infinity, and when $z \geq 1$, (5.64) goes to infinity.

Expanding (5.59), we have

$$R(\rho, \mu, \beta) = R_1(\rho, \mu, \beta) + R_2(\rho, \mu, \beta) + R_3(\rho, \mu, \beta), \quad (5.65)$$

where

$$R_1(\rho, \mu, \beta) = -\log e \left( \frac{F(x_1, z_1)}{4x_1} - \frac{F(x_2, z_2)}{4x_2} \right), \quad (5.66)$$

$$R_2(\rho, \mu, \beta) = z_1 \log(1 + x_1 - \frac{1}{4} F(x_1, z_1)) - z_2 \log(1 + x_2 - \frac{1}{4} F(x_2, z_2)), \quad (5.67)$$
$R_1(\rho, \mu, \beta) = \log(1 + x_1 z_1 - \frac{1}{4} F(x_1, z_1)) - \log(1 + x_2 z_2 - \frac{1}{4} F(x_2, z_2)),$  \hspace{1cm} (5.68)

with $x_i = \frac{\rho(1 + \frac{1}{\mu})}{\beta(L + 1)}$, $z_i = \beta(L + 1)$, $x_2 = \frac{\rho(1/\mu)}{\beta L}$, and $z_2 = \beta L$.

If $\beta > \frac{1}{L}$, then $z_1 > 1$, and $z_2 > 1$. By (5.62) we have

$$\lim_{\rho \to \infty} R_1(\rho, \mu, \beta) = 0.$$ \hspace{1cm} (5.69)

By (5.63) we have

$$\lim_{x \to \infty} R_2(\rho, \mu, \beta) = \beta(L + 1) \log \frac{1}{1 - \frac{1}{\beta(L + 1)}} - \beta L \log \frac{1}{1 - \frac{1}{\beta L}}.$$ \hspace{1cm} (5.70)

$\lim_{\rho \to \infty} R_1(\rho, \mu, \beta)$ is in the form of $\log \frac{\infty}{\infty}$. By L'Hopital's rule, we have

$$\lim_{\rho \to \infty} \frac{1 + x_1 z_1 - \frac{1}{4} F(x_1, z_1)}{1 + x_2 z_2 - \frac{1}{4} F(x_2, z_2)} = \lim_{\rho \to \infty} \frac{\frac{\partial}{\partial \rho} (1 + x_1 z_1 - \frac{1}{4} F(x_1, z_1))}{\frac{\partial}{\partial \rho} (1 + x_2 z_2 - \frac{1}{4} F(x_2, z_2))} = \frac{(1 + \frac{1}{\mu})}{\mu} \frac{z_1}{z_1 - 1} = \frac{1 \frac{1}{\mu}}{z_2} \frac{z_2 - 1}{z_2 - 1},$$ \hspace{1cm} (5.71)

where we use the fact that

$$\lim_{x_1 \to \infty} \frac{\partial}{\partial x_1} (1 + x_1 z_1 - \frac{1}{4} F(x_1, z_1)) = z_1 - 1$$ \hspace{1cm} (5.72)

when $z_1 > 1$. Equation (5.60) follows after simple calculation.

If $\beta \leq \frac{1}{L + 1}$, then $z_1 \leq 1$, and $z_2 < 1$. We have
\[
\lim_{\rho \to \infty} R_z(\rho, \mu, \beta) = (z_1 - z_2) \log(1 + x_1 - \frac{1}{4} F(x_1, z_1)) + z_2 \log \left( \frac{1 + x_1 - \frac{1}{4} F(x_1, z_1)}{1 + x_2 - \frac{1}{4} F(x_2, z_2)} \right).
\] (5.73)

Similarly,

\[
\lim_{\rho \to \infty} \frac{1 + x_1 - \frac{1}{4} F(x_1, z_1)}{1 + x_2 - \frac{1}{4} F(x_2, z_2)} = \lim_{\rho \to \infty} \frac{\partial}{\partial \rho} \frac{1 + x_1 - \frac{1}{4} F(x_1, z_1)}{1 + x_2 - \frac{1}{4} F(x_2, z_2)}
\]

\[
= \frac{(1 + 1/\mu)(1 - z_1)}{z_1} - \frac{1/\mu}{z_2} (1 - z_2),
\] (5.74)

where we use the fact that

\[
\lim_{x_1 \to \infty} \frac{\partial}{\partial x_1} (1 + x_1 - \frac{1}{4} F(x_1, z_1)) = 1 - z_1
\] (5.75)

when \( z_1 \leq 1 \). As \( z_1 > z_2 \), by (5.63) we have

\[
\lim_{\rho \to \infty} R_z(\rho, \mu, \beta) = \infty,
\] (5.76)

and (5.61) follows.
Chapter 6

Turbo Multiuser Detection for DSL Communications

6.1 Introduction

In this chapter, we apply some of the receiver techniques discussed in the preceding chapters to wireline communications. In particular, the turbo multiuser detection techniques are applied to digital subscriber line (DSL) wireline communications to effectively combat crosstalk, with the influence of impulse noise taken into consideration.

DSL technology provides transport of high-bit-rate digital information over telephone subscriber lines. Various DSL techniques (Basic Rate integrated service digital network (ISDN), high-rate DSL (HDSL), asymmetric DSL (ADSL), and very-high-rate DSL (VDSL)) involving sophisticated digital transmission schemes and extensive signal processing have recently become practical due to advances in microelectronics. The latest in DSL technology is VDSL, which provides tens of megabits per second to those customers who desire broadband entertainment or data services. At such high rates, signals on twisted pairs can be reliably transmitted at most to a few thousand feet. Thus, VDSL will primarily be used for loops fed from an optical network unit (ONU) or a central office (CO) to a customer premises, i.e., it addresses the so-called “last mile” problem. The modulation scheme for VDSL can either be multicarrier-based or single
carrier-based, typically discrete multitone (DMT) and carrierless amplitude/phase modulation (CAP)/quadrature amplitude modulation (QAM). The duplexing methods can be either time-division duplex (TDD) or frequency-division duplex (FDD) [23], [102].

Intersymbol interference (ISI) is one of the major obstacles to high-data-rate, bandwidth-efficient communications. Multicarrier modulation (MCM), following Shannon’s optimum transmission suggestion, achieves the highest performance in channels with ISI. DMT is a particular form of MCM that has been found to be well suited for DSL application and is adopted in ANSI T1.413 ADSL standards. With this approach, a channel is divided into many independent ISI-free subchannels in the frequency domain, and power and bits are allocated adaptively according to the channel characteristics [22], [102]. The advantages of using DMT for VDSL include optimality for data transmission, adaptivity to changing environments and flexibility in bandwidth management.

Typical phone lines that carry VDSL signals are 24- or 26-gauge unshielded twisted pairs (UTP). Multiple telephone pairs may share the same cable. Normally VDSL signals occupy from 300 kHz to 30 MHz of the twisted-pair bandwidth and are separated from plain old telephone system (POTS)/ISDN signals by splitter devices. Noise on phone lines normally occurs because of imperfect balance of the twisted pair. There are many types of noises that couple through imperfect balance into phone lines, the most common of which are crosstalk noise, radio noise and impulse noise. Crosstalk is caused by electromagnetic radiation of other phone lines in close proximity, in practice within the same cable. Such coupling increases with frequency and can be caused by signals traveling in the opposite direction, called near-end crosstalk (NEXT), and by signals
traveling in the same direction, called far-end crosstalk (FEXT). Radio noise is the remnant of wireless transmission signals coupling into phone lines, particularly AM radio broadcasts and amateur (HAM) operator transmissions. Impulse noise is a nonstationary crosstalk from temporary electromagnetic events (such as the ringing of phones on lines sharing the same binder, and atmospheric electrical surges) that can be narrowband or wideband and that occurs randomly. Impulse noises can be tens of millivolts in amplitude and can last as long as hundreds of microseconds.

This chapter is organized as follows. In Section 6.2 the DSL communication system model is presented. In Section 6.3 multiuser detection techniques are applied on the DSL system to combat the crosstalk. Robust multiuser detection techniques to jointly mitigate the crosstalk and the impulse noise are also discussed. In Section 6.4 Turbo multiuser detection for coded DSL system is presented. Numerical results are given in Section 6.5. Section 6.6 summarizes the chapter.

### 6.2 DSL System Model

We consider a (possibly encoded) DMT system with crosstalk as shown in Fig. 6.1. The information bits $d$ are first encoded into coded bits $b$ with a standard binary convolutional encoder with code rate $R$. A code-bit interleaver is used to decorrelate the noise on the coded bits at the input of the channel decoder. The interleaved bits are optimally allocated to $\bar{N}$ subchannels and mapped to QAM signals of various constellation sizes. Then the conjugate-symmetric vector of length $N = 2\bar{N}$ is transformed using the inverse
Fig. 6.1 VDSL DMT System Configuration: (a) Transmitter; (b) Channel; (c) Receiver
fast Fourier transform (IFFT) to get a real time-domain vector. After parallel-to-serial and digital-to-analog conversion, the DMT VDSL signal $x(t)$ is transmitted into the channel, where it is corrupted by additive coupled crosstalk signals and background noise. At the receiver end, after analog-to-digital and serial-to-parallel conversion, the received signal $r(t)$ is transformed back to the frequency domain using an FFT, where it can be written as

$$Y_i = H_i \cdot X_i + \sum_{m=2}^{M} F_{i,m} \cdot C_{i,m} + E_i, \quad i = 1,\ldots,N,$$  \hspace{1cm} (6.1)

where for the $i$th subchannel, $H_i$ is the channel gain, $X_i$ is the transmitted (complex) DMT symbol, $C_{i,m}$ is the $m$th crosstalk signal, $m = 2, \ldots, M$, $F_{i,m}$ is the corresponding crosstalk coupling function, and $E_i$ is the background noise. Output values of the FFT are fed into the demodulator and decoder for further processing. Note that when the VDSL signal and crosstalk signals are asynchronous (almost always the case), the crosstalk coupling functions can vary in time. However, our model considers only one DMT block at a time so we omit the time index for simplicity. In practice, the crosstalk coupling functions can be estimated for each block to allow the application of our models.

Impulse noise is a severe impairment to DSL transmission, especially after long loop attenuation (at a residential location) and in high frequencies (where the DSL signal is more severely attenuated). However, the area of impulse noise modeling remains unsettled. Cook presented an analytical model in [24]. The ADSL standard, however, uses stored representative impulse waveforms, which are measured empirically. Valenti
et al. collected impulse noise and background noise data on ADSL loops at New Jersey residences and conducted analysis on the data in three ways: as power and energy spectral densities, as probability density functions (pdf) of the time waveform voltage amplitudes, and as impulse arrival and inter-arrival time statistics [62], [109], [110]. So far there are no such models for impulse noise in VDSL, but similar results are anticipated [111]. Our key observation from these analyses is: there are significant impulse spikes in the PSD of the measured wideband noise, which is otherwise essentially flat. To model this behavior of the impulse noise we use a two-term Gaussian mixture model in the frequency domain. The first-order probability density function of this noise model has the form

\[
(1 - \varepsilon)N(0, \sigma^2) + \varepsilon N(0, \kappa \sigma^2),
\]

with \( \sigma > 0, \ 0 \leq \varepsilon \leq 1, \) and \( \kappa \geq 1. \) Here, the \( N(0, \sigma^2) \) term represents the nominal background noise (Gaussian with zero mean and variance \( \sigma^2 \)), and the \( N(0, \kappa \sigma^2) \) term represents an impulse component (Gaussian with zero mean and variance \( \kappa \sigma^2 \)), with \( \varepsilon \) representing the probability that impulses occur [124]. It is assumed that noise samples in disjoint frequency bins are independent.

### 6.3 Multiuser Detection for DSL

While the radio noise problem can be solved or at least alleviated by restricting the VDSL transmission in radio bands, crosstalk and impulse noise are two principal sources of degradation in VDSL transmission systems. The traditional single-user detector (SUD)
for such systems merges crosstalk into the background noise, which is assumed to be white and Gaussian. Actually, crosstalk is the result of the sum of several filtered discrete data signals. Its distribution deviates from Gaussian, and its power spectral density (PSD) is also significantly greater than that of background added white Gaussian noise (AWGN). Recent research has explored the nature of crosstalk signals and has shown the potential benefits of multiuser detection for VDSL signals with strong crosstalkers [16], [17], [21], [140]. In DSL transmission impulse noise is typically combated with interleaved forward error correction (FEC). However, recent survey data indicates that a significant minority of impulse noise events are longer than the maximum error correcting capacities of the default interleaved FEC provided within current ANSI standards [28], [29]. Thus, it is of interest to consider signal processing methods that can jointly mitigate crosstalk and impulsive noise.

Note that the crosstalk signals in DSL transmission are of various types and cannot be represented under a uniform framework, to the best of the author’s knowledge. In our application of MUD to signal detection in DSL systems, we deal mainly with NEXT of other types (in contrast to the self NEXT coming from the phone lines carrying the same VDSL service) for the following reason. FEXT experiences the same line attenuation as the desired signal while NEXT does not, which makes NEXT the most detrimental type of interference, especially at high frequencies. Self NEXT can be largely alleviated by duplexing methods that separate the upstream and downstream data in time or frequency. Therefore, the other-type NEXT provides the best opportunity for performance gain. Nevertheless, although we consider the other-type NEXT, multiuser detection is a valid
technique for mitigation of crosstalk of all types, albeit modifications of the techniques proposed here may be necessary for each specific situation.

### 6.3.1 Maximum Likelihood Multiuser Detection

Let us first consider the detection problem for the data model given in (6.1) for the case of Gaussian ambient noise. The traditional single user detector demodulates QAM symbols tone-by-tone independently. On the other hand, joint maximum-likelihood detection (ML-MUD) of both VDSL and crosstalk signals selects a set of $N$ inputs $\{X_i\}$ and the crosstalk sequences $C_m^{(l)} = \{C_{1,m}^{(l)}, C_{2,m}^{(l)}, \ldots, C_{N,m}^{(l)}\}$, $m = 2, \ldots, M$, to satisfy

$$X_i = \arg\min_{\{X_i\}, \{C_m^{(l)}\}} \sum_{i=1}^{N} \left| Y_i - H_i \cdot X_i - \sum_{m=2}^{M} F_{i,m} \cdot C_{i,m}^{(l)} \right|^2, \quad i = 1, \ldots, N, \quad (6.3)$$

where the minimization is taken over the DMT signal alphabet $\{X_i\}$ and all possible crosstalk sequences $\{C_{i,m}^{(l)}, l = 1, \ldots, |C_m|\}$, $m = 2, \ldots, M$, that can occur within the VDSL symbol period of interest. The size $|C_m|$, $m = 2, \ldots, M$, of the set of all possible crosstalk sequences can be large but is always finite when all the crosstalkers are digital signals or are derived from digital signals.

### 6.3.2 Interference Cancellation Multiuser Detection

Just as its counterpart in wireless CDMA does, the maximum-likelihood multiuser detector achieves optimum performance but suffers from very high complexity. A full search in the input domain requires approximately $N |C| |M|$ squared-error computations,
where $\bar{N}$ is the number of subchannels, $|C| = \prod_{m=2}^{M} |C_m|$ is the number of possible crosstalk sequences, and $|M|$ is the average size of the transmitted alphabet. In practice $\bar{N}$ and especially $|C|$ can be very large, introducing prohibitive computational complexity. The large number of possible crosstalk sequences also means an exponentially greater number of states, making dynamic programming inappropriate. Therefore, we need to consider a simplified receiver structure that maintains satisfactory performance while requiring far less computational complexity.

One lower-complexity approach to MUD is to employ a linear multiuser detection technique, such as decorrelating (zero forcing) or MMSE multiuser detection. However, unlike CDMA or space-division multiple-access (SDMA) where linear detection has been effective, there is no identifying signature such as the spreading code for CDMA or the steering vector for SDMA, to aid linear detection in VDSL. Moreover, the desired signals and crosstalk signals are often of different data format. An alternative approach that is better suited to this situation is to employ interference cancellation (IC), i.e., to attempt excision of the crosstalk from the received signal before applying traditional DMT VDSL signal detection.

Our interference cancellation multiuser detection (IC-MUD) scheme is described as follows (The type we examine here is the dominant near-end QAM-like crosstalk (e. g. [17])):

1. A hard decision is made on the VDSL signal in the frequency domain. This decision can be obtained through the received signal $\mathbf{Y} = (Y_1, \ldots, Y_{\bar{N}})^T$ directly or
through soft log-likelihood ratio (LLR) values from a soft-input soft-output (SISO) decoder (see 6.4).

(2). The VDSL signal \( \hat{\mathbf{X}} = (X_1, \ldots, X_R)^T \) are reconstructed based on these detected bits.

(3). The VDSL signal is subtracted and the known crosstalk coupling function is applied to get a frequency domain estimate of the entire crosstalk sequence \( \hat{\mathbf{C}}_m = (\hat{C}_{1,m}, \ldots, \hat{C}_{R,m})^T \) via

\[
\hat{\mathbf{C}}_m = \mathbf{F}_m^{-1} \circ (\mathbf{Y} - \mathbf{H} \circ \hat{\mathbf{X}} - \sum_{i< m} \mathbf{F}_i \circ \hat{\mathbf{C}}_i),
\]

where \( \mathbf{H} = (H_1, \ldots, H_R)^T \), \( \mathbf{F}_m = (F_{1,m}, \ldots, F_{R,m})^T \), \( \mathbf{F}_m^{-1} = (1/F_{1,m}, \ldots, 1/F_{R,m})^T \), \( \hat{\mathbf{C}}_i \) are formerly detected and recreated crosstalk signals, and “\( \circ \)” denotes Kronecker (elementwise) product.

(4). A time-domain sequence is obtained through the IFFT, \( \hat{\mathbf{c}}_m = \text{IFFT}(\hat{\mathbf{C}}_m) \), after which hard decisions are made on the crosstalk symbols and the crosstalk signal is recreated and transformed to the frequency domain to get \( \hat{\mathbf{C}}_m = (\hat{C}_{1,m}, \ldots, \hat{C}_{R,m})^T \).

(5). The above process is repeated until all crosstalk signals \( \hat{\mathbf{C}} = [\hat{C}_m, m = 2, \ldots, M] \) are estimated and reconstructed.

(6). Finally, SUD is used for DMT signal detection, i.e.,

\[
X_i = \arg \{ \min_{\{\hat{X}_i\}} \left| \frac{Y_i - H_i \cdot X_i - \sum_{m=2}^{M} F_{i,m} \cdot \hat{C}_{i,m}}{2} \right|^2 \}, \quad i = 1, \ldots, N. \quad (6.5)
\]
6.3.3 Robust Multiuser Detection with Impulse Noise

For white Gaussian noise, maximum likelihood detection is the same as least-squares (LS) curve fitting, as can be seen from (6.3). It is well known from the classic work of Tukey [108] that least-squares estimates are very sensitive to the tail behavior of the probability density of measurement errors (represented here by the additive noise). Its performance depends significantly on the Gaussian assumption, and even a slight deviation of the noise density from the Gaussian distribution can, in principle, cause a substantial degradation of the LS estimate. The LS estimate corresponding to (6.3) can be robustified by using the class of M-estimators proposed by Huber [58]. Instead of minimizing over a sum of squared residuals, Huber proposed to use a less rapidly increasing penalty function $\rho$ so as to alleviate the effect of the impulses.

$$X_i = \arg\{\min_{\{X_j, (C_{i,m})\}} \sum_{i=1}^{5} \rho\left| Y_i - H_{i} \cdot X_i - \sum_{m=2}^{M} F_{i,m} \cdot C_{i,m}^{(i)} \right|\}, \quad i = 1, \ldots, N. \quad (6.6)$$

The usual requirements for the penalty function and its derivative $\psi = \rho'$ are:

1. $\rho$ is sub-quadratic function for large values of residuals, in order to de-emphasize the error caused by noise "outliers" (in this case caused by impulse noise);

2. $\psi$ is bounded and continuous;

3. $\psi(x) \approx kx$ for small $x$, so as to achieve high efficiency in the Gaussian case;

4. $E\{\psi(N_j)\} = 0$ to get a consistent estimate; and for symmetric noise densities $\psi$ is usually odd symmetric.

A good choice for Gaussian mixture noise is the Huber penalty function $\rho$ shown in Fig. 6.2 together with its derivative $\psi$. These functions are given explicitly by
\[ \rho(x) = \begin{cases} \frac{x^2}{2\sigma^2} & |x| \leq k\sigma^2 \\ k |x| - \frac{k^2\sigma^2}{2} & |x| > k\sigma^2 \end{cases} \quad (6.7) \]

and

\[ \psi(x) = \begin{cases} \frac{x}{\sigma^2} & |x| \leq k\sigma^2 \\ k \text{sgn}(x) & |x| > k\sigma, \end{cases} \quad (6.8) \]

where \( k, \sigma, \) and \( \varepsilon \) (see (6.2)) are related by

\[ \frac{\phi(k\sigma)}{k\sigma} - Q(k\sigma) = \frac{\varepsilon}{2(1 - \varepsilon)}, \quad (6.9) \]

with \( \phi(x) \triangleq \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \) and \( Q(t) \triangleq \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{x^2}{2}} dx \) (see [58]).

Fig. 6.2 Huber penalty function and its derivative for the Gaussian mixture model

\( \varepsilon = 0.1, \; \kappa = 100, \; \sigma^2 = 1, \; k = 1.14 \)
Although we only talk about the robustification of ML MUD here, the idea can be straightforwardly applied to other sub-optimal MUD schemes such as IC MUD. The resulting DMT VDSL detector will be called the robust maximum likelihood multiuser detection receiver (ML-MUD-R) and the robust interference cancellation multiuser detector (IC-MUD-R), respectively.

### 6.3.4 Importance Sampling Techniques for Intensive Simulations

In the DSL environment, BER values as low as $10^{-7}$ are often required. For Monte Carlo (MC) simulation, approximately $10/P_e$ simulation trials are required to have a 95 percent confidence interval of $[2P_e/5, 8P_e/5]$ [76]. To alleviate this computational burden, importance sampling (IS) techniques are adopted [76], [77], [92]. The basic idea of importance sampling is to bias the probability density function from which the data are generated so that errors of detection are more likely to happen, then weight each error such that an unbiased BER estimate is obtained. Assume an error occurs when the received data $R$ falls within some region $Z$. Then the BER is given by

$$P_e = \int 1_Z(r)f_R(r)dr,$$

where $1_Z(\cdot)$ is the indicator function over $Z$ and $f_R(\cdot)$ is the pdf of $R$.

The MC estimator of $P_e$ is given by

$$\hat{P}_{MC} = \frac{1}{M} \sum_{i=1}^{M} 1_Z(R_i),$$

(6.11)
where $M$ is the number of trials of the simulation and the $R_i$’s denote data samples. When the data samples are i.i.d., $\hat{P}_{MC}$ is an unbiased estimator with variance

$$\text{var}(\hat{P}_{MC}) = \frac{P_e (1 - P_e)}{M}. \quad (6.12)$$

The IS estimator of $P_e$ is given by

$$\hat{P}_{IS} = \frac{1}{M} \sum_{i=1}^{M} I_{Z} (R_i^*) W(R_i^*) \quad (6.13)$$

with

$$W (r) = \frac{f_{R_e} (r)}{f_{R_e'} (r)}, \quad (6.14)$$

where $R_i^*$ is the $i$th data sample from biased density $f_{R_e'} (\cdot)$ and $W (\cdot)$ is the weight function. If the newly generated data are i.i.d., $\hat{P}_{IS}$ is an unbiased estimator with variance

$$\text{var}(\hat{P}_{IS}) = \frac{\bar{W} - P_e^2}{M}, \quad (6.15)$$

where $\bar{W}$ is defined as

$$\bar{W} = \int_{Z} W (r) f_{R_e} (r) dr. \quad (6.16)$$

When $f_{R_e'} (\cdot)$ is appropriately selected, the variance of the IS estimator will be far less than that of the MC estimator. Thus the number of trials needed for a given estimator variance is greatly reduced for the IS estimator compared to the MC estimator. The optimal bias distribution is given by
\[ f_{R_{\text{opt}}}(r) = \frac{1}{P_e} f_p(r), \quad (6.17) \]

which achieves zero estimation variance but is degenerate since it requires the knowledge of \( P_e \). A widely used method of designing suboptimal \( f_{R'}(\cdot) \) is mean translation (MT).

This class of biased density functions is of the form

\[ f_{R'}(r^*) = f_p(r^* - T), \quad (6.18) \]

where \( T \) is chosen to be the mode (at which maximum value of a pdf is achieved) of \( f_{R_{\text{opt}}}(\cdot) \).

For the multiuser communication system of (6.1), let \( \rho_i = (X_i, C_{i,2}, \ldots, C_{i,M}) \), impose the restriction \( f_{E_i}(\cdot) = f_{\rho_i}(\cdot) \) and conditionally shift the mean of the noise

\[ f_{E_i, \rho_i}(E^*_i) = f_{E_i}(E^*_i + m(\rho_i^*)) = f_{E_i}(E^*_i + H_i \cdot X^*_i + \sum_{m=2}^{M} F_{i,m} \cdot C^*_i), \quad (6.19) \]

the IS estimator of BER is then given by

\[ \hat{P}_{\text{IS}} = \frac{1}{M} \sum_{i=1}^{M} I(\| \hat{X}^*_i - X^*_i \| W(\rho_i^*, E^*_i)) = \frac{1}{M} \sum_{i=1}^{M} I(\| \hat{X}^*_i - X^*_i \|) \frac{f_{E_i}(E^*_i)}{f_{E_i, \rho_i}(E^*_i)}, \quad (6.20) \]

where we assume the independence of \( \rho_i \) and \( E_i \); \( \hat{X}^*_i \) is the detected data of \( X^*_i \) with the original decision rule, and

\[ I(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0. \end{cases} \quad (6.21) \]

When the near-far problem occurs, i.e.,
we need to adjust the IS error estimator as follows:

\[
\hat{P}_{IS} = \frac{1}{M} \sum_{i=1}^{M} \left( I(\hat{X}_i^* - X_i^*) + (1 - I(\hat{X}_i^* - X_i^*)) \left(1 - \frac{f_{E_i}(E_i^*)}{f_{E_i|\psi_i}(E_i^*)}\right) \right). \tag{6.23}
\]

Note that in this situation, it is better using the IS technique to count correct detections, which happen with small probability. According to (6.19) and (6.20), we have

\[
\hat{P}_{IS-correct} = \frac{1}{M} \sum_{i=1}^{M} (1 - I(\hat{X}_i^* - X_i^*)) \frac{f_{E_i}(E_i^*)}{f_{E_i|\psi_i}(E_i^*)}, \tag{6.24}
\]

The quantity of (6.23) then follows through \( \hat{P}_{IS} = 1 - \hat{P}_{IS-correct} \).

In our simulations, the IS technique is uniformly better than the MC technique. It achieves great variance reduction for optimal detection (maximum likelihood) methods and substantial gains for others.

### 6.4 Turbo Multiuser Detection for Coded DSL

Coding is a common way to reduce the gap in channel capacity experienced by uncoded systems. A concatenated coding scheme consisting of an inner trellis code (a 4-D Wei’s code) and an outer Reed-Solomon (RS) code was proposed for ADSL DMT systems to provide a 5 dB coding gain at bit error rate (BER) \(10^{-7}\) without bandwidth expansion [134], [135]. There are two problems with this approach. First, since the constellation size varies from tone to tone, a time-varying trellis-coded modulation (TCM) encoder is
required. Second, further improvement is very difficult from a practical implementation perspective because of the complexity of Viterbi decoding for multidimensional TCM. Recently powerful turbo coding has been proposed for DMT systems [8], [47], [66], [122]. One typical turbo code is used to code across all subchannels. The coded bits are then interleaved and allocated to various tones for quadrature amplitude modulation. Thus, a single standard binary decoder can be employed at the receiver and further improvement in turbo coding is easily incorporated. A coding gain of 6.0 dB for bandwidth efficiency of 2 bits/s/Hz and 4.1 dB for 3 bits/s/Hz was reported for a channel with severe ISI at BER $10^{-5}$ [66].

Here we consider the application of turbo multiuser detection in a coded DMT VDSL system to combat crosstalk and to obtain substantial coding gain. We also consider the effects of impulse noise, which has been found to greatly impact the performance of our proposed receiver, and an erasure decoding technique is proposed as a remedy.

![Turbo structure for iterative demodulation and decoding](image)

**Fig. 6.3 Turbo structure for iterative demodulation and decoding**

Figure 6.3 shows the turbo structure for iterative demodulation and decoding. It consists of two stages: a soft metric calculator (the demodulation stage) and a SISO channel decoder (the decoding stage). The two stages are separated by an interleaver and a de-
interleaver. The crosstalk signals are first detected via a multiuser detection technique, discussed in the last subsection. Then a channel log-likelihood ratio for the $k$th bit carried by the $i$th subchannel symbol is calculated as follows:

$$
\Lambda_i(b_{k,i}) = \log \frac{P(b_{k,i} = 1 | \{r(t)\})}{P(b_{k,i} = -1 | \{r(t)\})}, \quad (6.25)
$$

where $\{r(t)\}$ is the received waveform as shown in Fig. 6.1(c). Using Bayes’ formula and discarding the common term $p(\{r(t)\})$, (6.25) can be written as

$$
\Lambda_i(b_{k,i}) = \log \frac{p(\{r(t)\} | b_{k,i} = 1)}{p(\{r(t)\} | b_{k,i} = -1)} + \log \frac{P(b_{k,i} = 1)}{P(b_{k,i} = -1)} + \lambda_i^p(b_{k,i}), \quad (6.26)
$$

where the second term $\lambda_i^p(b_{k,i})$ represents the a priori LLR delivered from the decoding stage in the previous iteration. For the first iteration, this term is set to zero if we assume equally likely coded bits. The first term $\lambda_i(b_{k,i})$, denoting the extrinsic information obtained from the demodulation stage about the bit $b_{k,i}$, is then de-interleaved and sent to the channel decoder as its a priori information. Similarly, the SISO channel decoder computes the a posteriori LLR of each code bit and then excludes the influence of a priori knowledge to get extrinsic information from the decoding stage about the bit $b_j$ (the de-interleaved version of $b_{k,i}$, alternatively the coded bits before the interleaver in Fig. 6.1(a)) as follows:

$$
\lambda_i^p(b_j) = \Lambda_i(b_j) - \lambda_i^p(b_j) = \log \frac{P(b_j = 1 | \text{decoding})}{P(b_j = -1 | \text{decoding})} - \lambda_i^p(b_j). \quad (6.27)
$$
Again, this extrinsic information is interleaved and fed back to the demodulation stage as *a priori* knowledge for the next iteration. At the last iteration, the SISO decoder also computes the *a posteriori* LLRs for information bits, which are used to make final decisions. More details on this turbo decoding process is given in the following.

### 6.4.1 Turbo Decoding for Coded DMT System

In contrast to trellis-coded modulation where coding and modulation are considered jointly to obtain an optimal signal constellation in the sense of maximum Euclidean distance in the signal space, the iterative decoding process we propose separates demodulation and channel decoding stages for ease of implementation. The greatest benefits of this turbo process are the very large increase in code memory due to the interleaver, and almost no loss of optimality due to the iterative exchange of extrinsic soft information.

Assuming the independence of the background noise between subchannels and between the two dimensions of each subchannel, the soft metric delivered by the demodulation stage (see (6.26)) can be expressed as

\[
\lambda_i(b_{k,i}) = \log \frac{p(r(t) \mid b_{k,i} = 1)}{p(r(t) \mid b_{k,i} = -1)} = \log \frac{p(Y_i \mid b_{k,i} = 1)}{p(Y_i \mid b_{k,i} = -1)} = \log \frac{p(Y_i, b_{k,i} = 1)P(b_{k,i} = -1)}{p(Y_i, b_{k,i} = -1)P(b_{k,i} = 1)}
\]
Here the second equality holds as $Y = (Y_1, \ldots, Y_T)^T$ is a sufficient statistic for $\{r(t)\}$ and has independent components. The third equality comes from Bayes’ formula. The $L_i$ different bits $b_i = \{b_{i,j}, \ldots, b_{i,L_i}\}$ assigned to a symbol $X_i$ can be modeled as being independent due to the interleaving effect, which leads to the last equality, where the summations are over all possible DMT symbol realizations with the indicated conditions. Note that, to be exact, the marginal distribution for the crosstalk signals should be incorporated in the last equality of (6.28). But this unrealistic approach would hinder the implementation of the proposed algorithm in practice. Instead, we use the MUD schemes discussed in 6.3 to detect the crosstalk signals first and subtract them from the received signal. For ML-MUD, the crosstalk signals are detected and subtracted from the received signal before iteration begins. As we can see, ML-MUD makes a thorough search over the product space spanned by the desired VDSL DMT signal and the crosstalk signals, so the refined estimates from the decoding stage for the data bits would not help here. However, for IC-MUD, after each iteration, the LLRs from the SISO decoder are used to help better detect the crosstalk signals, and ultimately obtain a better estimate of the code.
bits at the soft metric calculator (see step 1 of the IC-MUD scheme in 6.3.2). For the last expression in (6.28), we have the formula

$$P(b_{k,i}) = \frac{1}{2} \left[ 1 + b_{k,i} \tanh(\lambda_{2}^{p}(b_{k,i})) \right],$$

(6.29)

where $\lambda_{2}^{p}(b_{k,i})$ represents the corresponding a priori LLR delivered from the channel decoding stage in the previous iteration. Invoking the Gaussian assumption on the background noise, we can also write

$$p(Y_{i} | X_{i}, \{\hat{C}_{l,m}\}_{m=2}^{M}) = \frac{1}{2\pi\sigma^{2}} e^{-\frac{Y_{i} - H_{i}X_{i} - \sum_{m=2}^{M} E_{i,m} \hat{C}_{l,m}}{2\sigma^{2}}},$$

(6.30)

where $\sigma^{2}$ is the noise variance per dimension. Due to the independence of the in-phase and quadrature data, the above formulas (6.28) through (6.30) can be treated for each dimension separately.

For the channel decoding stage, algorithms discussed in 4.3.2 can be used and details are ignored here.

### 6.5 Numerical Results

#### 6.5.1 Robust Multiuser Detection with Impulse Noise

In this subsection we examine the behavior and the performance of the proposed multiuser detection receivers for DMT-VDAL signals with crosstalk and impulse noise
via computer simulations. Bit-error-rate (BER) is adopted as the performance measure with respect to the average signal-to-noise ratio (SNR), which is defined as

$$\text{SNR} = \frac{\sum_{i=1}^{N_i} |H_i \cdot X_i|^2}{\sum_{i=1}^{N_i} |E_i|^2}.$$  

(6.31)

In the simulation, the DMT VDSL signal is assumed to occupy 0-25.6 MHz with 256 subchannels in an frequency-division multiplexed (FDM) design. The symbol rate for each VDSL subchannel is 100 kHz. In each subchannel, 2 bits are assigned so the signals are 4-QAM. No bit allocation algorithms are used here, although extension to this case is straightforward. The transfer function of the DMT VDSL signal is simulated by

$$H(e^{j\omega}) = \frac{2 - 4e^{-j\omega} - 2e^{-2j\omega}}{0.965 - 1.50e^{-j\omega} + 0.539e^{-2j\omega}} \times 10^{-3}.$$  

(6.32)

We assume one NEXT crosstalk signal with a known coupling function given as

$$F(e^{j\omega}) = K \cdot \omega^{3/4},$$  

(6.33)

where $K$ is a constant used to adjust the PSD of the crosstalk signal. These settings are made to roughly approach the PSD shapes indicated in [10]. We assume that these transfer functions stay fixed for the whole simulation interval, which is reasonable for wireline communications environments. The crosstalk signal is BPSK modulated on 8 MHz carrier frequency with a 1M symbol-per-second rate. Such a situation would arise, for example, due to the coexistence of home-phone LANs and asymmetric DMT VDSL signals in the same cable in the customer premises [16], [17], [18]. Thus, there are $2^{10}$ possible crosstalk sequences in one VDSL symbol. This number is chosen for simulation simplicity. In reality, this number could be much larger. The impulse noise is assumed to
have parameters $\varepsilon = 0.1$ and $\kappa = 100$, which means the impulse spike is 20 dB higher than the background noise floor. The average PSD levels of the crosstalk signal and background noise floor are fixed while that of the desired signal is varied, corresponding to different line lengths (the signal attenuation is increasing with the line length). In our simulation, the average PSD of the crosstalk is 27 dB higher than that of the background noise floor and the peak PSD of the crosstalk is 40 dB higher. These settings seem to agree roughly with empirical measurements.

![Graph showing BER versus SNR](image)

**Fig. 6.4** Bit error rate (BER) versus signal-to-noise ratio (SNR) for different detectors (x-mark: SUD, circle: IC-MUD, diamond: ML-MUD, dashed: single user lower bound)

Figure 6.4 shows the performance of various detectors for DMT VDSL systems with one crosstalker and impulse noise. As we can see, there is a significant gap between the performance of the traditional single user detector and the single user lower bound (corresponding to a crosstalk-free channel), indicating the ineffectiveness of the single-user detector. While the maximum likelihood multiuser detector essentially achieves the
single user lower bound, it suffers from prohibitive complexity. The interference
cancellation multiuser detector offers a favorable performance and complexity tradeoff
compared with the single-user and ML multiuser detectors.

![Graph showing BER versus SNR for different detectors](image)

**Fig. 6.5** Bit error rate (BER) versus signal-to-noise ratio (SNR) for different
detectors (x-mark: SUD, circle: IC-MUD, plus: IC-MUD-R, diamond: ML-MUD,

Figure 6.5 shows the performance of the M-estimator-based robust detectors in the
crosstalk and impulse noise environment. While there is not much difference between the
ML multiuser detector and its robust version, both of which approximate the single user
lower bound, there is significant improvement for the robust interference cancellation
multiuser detection compared with its Gaussian-based counterpart. The crosstalk
detection errors are \(3.42 \times 10^{-4}\) for IC-MUD, \(6.10 \times 10^{-5}\) for IC-MUD-R and almost 0 for
ML-MUD and ML-MUD-R. It should be noted that the expected improvement from
using M-estimators is due to better estimation of the crosstalk signals in the DMT VDSL
case. The desired DMT VDSL signals in different subchannels are independent while the crosstalk signals are correlated in the frequency domain, which means that M-estimators are especially applicable to impulse-noise-contaminated DMT VDSL systems with crosstalk signals strongly correlated in the frequency domain.

6.5.2 Turbo Multiuser Detection

In this subsection we examine the behavior and the performance of the proposed turbo multiuser detection receivers for coded DMT-VDSL signals with crosstalk via computer simulations. The main results are with an AWGN channel, although impulse noise issues are also addressed briefly. BER is adopted as the performance measure with respect to the geometric SNR, which is defined as

$$\text{SNR}_{\text{geo}} = \Gamma \cdot \left[ \left( \prod_{i=1}^{\mathcal{N}} (1 + \frac{\text{SNR}_i}{\Gamma}) \right)^{1/\mathcal{N}} - 1 \right], \quad (6.34)$$

where \(\text{SNR}_i\) is the SNR on the \(i\)th subchannel, and \(\Gamma\) is the SNR gap to capacity [22] which is defined as

$$3 \cdot \Gamma = \frac{d_{\text{min}}^2}{4\sigma^2}, \quad (6.35)$$

where as before \(\sigma^2\) is the ambient noise variance per dimension, which is assumed the same for all subchannels, and \(d_{\text{min}}\) is the minimum Euclidean distance of the received signals, which is also assumed to be the same for all used subchannels. The rationale for achieving the same gap for all used subchannels is that the aggregate performance is approximately maximized if the BERs in all used subchannels are equal. Note that
where $d_i$ is constellation point distance for the $i$th subchannel. Assuming that the $i$th subchannel carries $c_i$ bits of information using $M = 2^c$ QAM modulation, we have

$$c_i = \log_2 \left( 1 + \frac{\text{SNR}_i}{\Gamma} \right).$$

Then the total number of bits transmitted in one DMT symbol is given by

$$c = \sum_i c_i = \sum_i \log_2 \left( 1 + \frac{\text{SNR}_i}{\Gamma} \right) = \bar{N} \cdot \log_2 \left( 1 + \frac{\text{SNR}_{geo}}{\Gamma} \right).$$

Thus, the geometric SNR allows a direct comparison of performance for all single and multicarrier systems with the same bit rate $\frac{c}{\bar{N}}$. The SNRs used in this simulation are measured as $\frac{E_b}{N_0}$, the energy per information bit divided by the one-sided noise power spectral density ($N_0 = 2\sigma^2$).

In the simulations, the DMT VDSL signal is assumed to occupy 0-25.6 MHz with 256 subchannels in a FDM design. The symbol rate for each VDSL subchannel is 100 kHz. A rate-1/2 convolutional code with constraint length 5 and generator polynomials $[23, 35]_8$ is used for channel coding. The number of coded bits per data frame is set as 1024, indicating an average bit rate of 512 bits per DMT block or 2 bits/s/Hz. A random interleaver of length 1024 is used for interleaving and de-interleaving. The coded DMT system is applied to a channel with severe ISI, the transfer function of which is taken from [66] to be

$$H(\omega) = 1 + 0.9 \cos(\omega T).$$
This channel is not necessarily typical of a VDSL line transfer function, however it serves as a useful example for illustration purposes. Campello’s margin-adaptive bit-loading algorithm [14] is used to allocate bits to DMT subchannels. We assume a square QAM constellation for simplicity, so the granularity of bit loading equals 2. Also the first two tones (up to 200 KHz) are not used for compatibility with POTS/ISDN service. Figure 6.6 gives the bit allocation used for our simulation. We see that typical constellations are 64-, 16- and 4-QAM. Subchannels not able to transmit 2 bits reliably are not used.

![Bit allocation for DMT subchannels](image)

Fig. 6.6 Bit allocation for DMT subchannels

We assume one NEXT crosstalk signal with a known coupling function. The crosstalk signal is binary phase-shift keying (BPSK) modulated, carried on a 12.8 MHz central frequency with a 0.8M symbol-per-second rate. The average PSD of the crosstalk
Fig. 6.7 Performance of the ML-MAP turbo multiuser receiver

Fig. 6.8 Performance of the IC-MAP turbo multiuser receiver
CHAPTER 6. TURBO MULTIUSER DETECTION FOR DSL COMMUNICATIONS

Fig. 6.9 Performance of the ML-SOVA turbo multiuser receiver

Fig. 6.10 Performance of the IC-SOVA turbo multiuser receiver
is 30 dB higher than that of the background noise floor, and the peak PSD of the crosstalk is 48 dB higher.

Figures 6.7 to 6.10 give, respectively, the performance for the first five iterations for four different receivers exploiting either ML or IC for MUD detection, and MAP or SOVA for channel decoding. These receivers are called ML-MAP, IC-MAP, ML-SOVA and IC-SOVA. From these figures, we see that turbo processing monotonically improves the system performance, with large gain obtained in a few iterations. We also observe that the IC-MUD scheme, incorporated into the turbo process, performs as well as the ML-MUD scheme does with much lower computational requirements. However, IC-MUD is sub-optimal when used to combat crosstalk signals alone, which we will see shortly. Finally, SOVA decoding performs almost as well as MAP, with only a few tenths of a dB loss, but with the advantage of much lower complexity.

To get an overall idea of how much gain is obtained through turbo multiuser detection of a coded DMT system, we compare the performance of the proposed schemes with their SUD, uncoded and non-iterative counterparts in Fig. 6.11. Here, SUD means application of traditional demodulation to the uncoded DMT system, while IC-MUD and ML-MUD mean application of the corresponding multiuser detection scheme on an uncoded system. Clearly, MUD greatly outperforms SUD. For BER greater than $10^{-3}$, IC-MUD performs identically with ML-MUD. But when the desired signal level increases, the performance of IC diminishes. (It is well known that IC performs worst when the powers of the different users are equal or comparable.) IC-MUD+VA and ML-MUD+VA refer to detection of a coded DMT VDSL signal in a non-iterative way: after multiuser detection, hard decisions are made on coded bits, then the Viterbi algorithm is
used for decoding. The coded DMT system has a smaller $d_{\text{min}}$ compared with its uncoded counterpart for the same $\frac{E_b}{N_0}$, but the coding gain more than offsets this disadvantage. We see that at BER $10^{-7}$, ML-MUD+VA provides 2.5 dB gain over ML-MUD. Again, the IC-MUD+VA scheme deteriorates for higher DMT signal levels. The ML-MUD+MAP and IC-MUD+SOVA denote the performance of our turbo iterative algorithms, ML-MAP and IC-SOVA, at the fifth iteration. At BER $10^{-7}$, we see that an additional 4.5 dB gain is achieved over ML-MUD+VA.

For completeness, we include the performance of SUD+MAP, which means the turbo iterative decoder (with MAP algorithm for channel decoding) without the MUD processing step (either ML-MUD or IC-MUD), at the fifth iteration. It can be seen from the figure that, without multiuser detection, the received signal is crosstalk-dominated.
and the turbo decoding process is useless. We have also examined the performance of the turbo iterative receiver when the IC-MUD is executed only once at the first iteration and found the performance is identical with our proposed receiver which incorporates the IC-MUD in the iterative process. In other words, in our settings one-time IC-MUD is enough. That is because in the BER range of interest, the power level of the desired VDSL signal is so small compared to that of the dominant NEXT, so the IC scheme has excellent performance here. Nonetheless, we incorporate the IC-MUD in the turbo process in our proposed scheme for generality and believe it will help in other situations.

In the following, we further address two related issues for our proposed iterative DMT receiver. For simplicity, we do not include crosstalk signals (thus no need for MUD in the SISO demodulation stage) and MAP is used in the SISO decoder. Usually in multilevel QAM symbol mapping, Gray coding is used so that a symbol detection error results in only one bit error. For our proposed iterative decoding process, we found that natural coding actually outperforms Gray coding. Figures 6.12 and 6.13 give the performance of the iterative DMT receiver with Gray coding and natural coding for first five iterations. It is shown that with Gray coding, the receiver is saturated after two iterations with performance worse than natural coding.

This phenomenon can be attributed to the reduced minimum Euclidian distance associated with Gray symbol mapping. Consider, for example, a one-dimension signal level of 16-QAM. Figures 6.14 and 6.15 show Gray coding and natural coding for a 4-PAM constellation. In our proposed scheme, the demodulation stage yields a soft metric for each coded bit based on the received noise-contaminated signal and a priori information of all other bits of the same DMT symbol from the decoding stage. Let us
Fig. 6.12 Performance of the iterative DMT receiver with Gray coding

Fig. 6.13 Performance of the iterative DMT receiver with natural coding
denote the associated bits for each symbol in Fig. 6.14 and Fig. 6.15 as “ab” and the
constellation point distance as $d$. If we assume the correctness of the soft information for
the other bit from the decoding stage, with Gray coding, the minimum distances for bit $a$
and $b$ are both $d$. Note that although when $b = 0$, the Euclidean distance for bit $a$, which is
the distance between symbols associated with “00” and “10”, is $3d$, this minimum
distance becomes $d$ when $b = 1$. This distance essentially determines the reliability of the
soft metric produced in the demodulation stage. In contrast, for natural coding, the
minimum distance is $d$ for bit $b$ and $2d$ for bit $a$. Similarly, for a 64-QAM constellation,
the minimum distance is $d$ for all bits with Gray coding, while it is $d$, $2d$, and $4d$ for
different bits respectively with natural coding. There may be other reasons for this
phenomenon, which deserve further study.

![Fig. 6.14 Gray coding for 4-PAM](image1)

![Fig. 6.15 Natural coding for 4-PAM](image2)

Finally, we would like to examine the performance of this proposed iterative DMT
receiver with impulse noise. We model the behavior of the impulse noise using a two-
term Gaussian mixture model in the frequency domain as in (6.2) with parameters
$\varepsilon = 0.01$ and $\kappa = 100$. Again we do not include crosstalk signals and MAP is used in the
SISO decoder. Figure 6.16 shows that the performance of the proposed receiver is greatly
degraded with impulse noise. The use of erasure decoding can remedy this. In the
demodulation stage, for those bits associated with impulse-contaminated symbols, instead
of calculating soft metrics for them, the \textit{a priori} information is used as a substitute, i.e.,
\begin{equation}
\lambda_i(b_{k,j}) = \lambda_{2}^u(b_{k,j}).
\end{equation}
(For the first iteration, these are set to zeros.) In DMT systems, the erasure positions where impulse spikes appear can possibly be detected in advance through pilot tones. In Fig. 6.17, we see that the performance of the proposed receiver experiences almost no performance loss with impulse noise with the aid of erasure decoding.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{impulse_noise_graph.png}
\caption{Performance of the iterative DMT receiver with impulse noise}
\end{figure}
6.6 Summary

In this chapter we have shown the potential benefits of multiuser detection for crosstalk mitigation in DMT VDSL systems subject to impulse noise. We see that ML-MUD can essentially eliminate crosstalk signals in DMT systems at a cost of high complexity. As a tradeoff, IC-MUD can significantly outperform SUD, with lower complexity than ML-MUD. We have also shown the effectiveness of the M-estimator in combating the impulse noise.

Further, a new coded DMT VDSL receiver structure using turbo multiuser detection has been proposed and has been shown to achieve an overall 7.0 dB gain over the uncoded optimum (maximum likelihood) receiver at BER $10^{-7}$ for a channel with severe...
ISI, AWGN, and one dominant crosstalk signal. The traditional single-user detection scheme produces extremely poor and totally unacceptable performance in our settings. Without multiuser detection, the decoding process turns out to be useless. The effect of impulse noise is detrimental to the proposed scheme but can be overcome through an erasure decoding technique. For our proposed scheme, we also see that natural coding is better than Gray coding.

In this chapter, we have assumed knowledge of the line transfer function and crosstalk coupling functions. In reality, however, channel identification is needed, and the effects of channel estimation error should be taken into consideration. This issue is of interest for further study. The problem of detecting impulse spike positions (for erasure decoding purposes) also deserves further study.
Conclusions and Perspectives

In this dissertation, both signal processing and information theoretic aspects of wireless cellular communications with antenna arrays have been studied. Specifically, we have investigated various transmit diversity and downlink beamforming techniques with power control for wireless cellular communications with transmit arrays, and iterative space-time multiuser detection techniques with receive arrays. Further, turbo space-time multiuser detection has been employed for wireless cellular MIMO communications. Then, for the multicell MIMO systems where co-channel interference is the dominating detrimental factor, various multiuser receivers have been proposed to dramatically improve the system performance. Spectral efficiencies of these receivers have been analyzed, together with their large-system asymptotic (non-random) expressions. The turbo multiuser detection techniques have also been applied to the DSL wireline communications to effectively combat the crosstalk, with the influence of impulse noise taken into consideration.

The contributions of this dissertation are summarized in the following.

Chapter 2: It has been shown that traffic type impacts algorithm choice. For circuit-switched downlink CDMA systems, Max SNR beamforming is the best choice (accommodates 12 to 42 more users than transmit diversity). For packet-switched
systems, Max SINR Beamforming has the best performance in terms of peak rate (10-14 dB more than transmit diversity); Max SNR beamforming is almost the best in terms of mean rate (3-4 dB more than transmit diversity), but beam steering and transmit diversity with sectorization are also good choices. It has also been found that sectorization greatly improves the system performance, both for the circuit and for the packet case. For transmit diversity, the gain through exploiting more antennas diminishes, while this is not the case for beamforming techniques, especially with feedback channel parameters.

**Chapter 3**: For batch iterative methods, it has been shown that fully exploiting diversities through space-time processing and multiuser detection offers substantial improvement over alternative processing methods, and that iterative implementation of these linear and nonlinear multiuser receivers realizes this substantial gain and approaches the optimum performance with reasonable complexity. Among these iterative implementations the SAGE space-time multiuser receiver outperforms the others. The complexity of the new SAGE detector with the traditional structure is no higher than the existing methods but with better performance and smoother convergence. The SAGE detector with a new structure retains its excellent performance but with greater adaptability, and its complexity is comparable to the existing methods. For sample-by-sample adaptive methods, it is shown that the data aided adaptive space-time multiuser receivers combine the functions of adaptive beamforming, RAKE combining and multiuser detection with no side information needed other than the timing and training sequences of the desired user. The alternative blind adaptive space-time multiuser receiver is based on the LCMV criterion and min-max parameter estimation. This detector is robustified against signature waveform mismatch with norm-constrained
techniques. LMS implementations of all of these adaptive ST MUD receivers have been given, and the convergence and excess-MSE issues have been discussed.

Chapter 4: Turbo space-time multiuser detection has been introduced for intracell wireless MIMO communications. Among various multiuser detection techniques examined to combat the intercell interference, linear MMSE MUD and successive interference cancellation have been shown to be feasible and effective. Based on these two multiuser detection schemes, one of which may outperform the other for different settings, an adaptive detection scheme has been developed, which together with a Turbo-BLAST structure offers substantial performance gain over the well known V-BLAST techniques with coding in this interference-limited cellular environment. The obtained multiuser capacity is excellent in high to medium SIR scenario. Nonetheless, numerical results also indicate that a further increase in system complexity, using base-station cooperation, could lead to further significant increases of the system capacity.

Chapter 5: The spectral efficiency of the multicell MIMO systems with several MUD detectors has been studied, among which are single-cell detectors, joint optimum detectors, group linear MMSE detectors, group MMSE successive cancellation detectors, and adaptive multiuser detectors. The large-system asymptotic (non-random) expressions for these spectral efficiencies are also explored. As applications of these theoretical bounds, we have verified the following observations of Chapter 4 about multicell MIMO systems. Group linear MMSE detection and group MMSE successive cancellation are two effective techniques to combat the co-channel interference, each of which may outperform the other for different settings. For full system load, multiuser detection across the cell is most useful in strong interference environments, offering substantial
gain over traditional single-cell detectors. However, there is still a substantial gap between the achievable capacity and the single cell upper bound in this case. Further, the conditions for non-interference-limited behavior of the group linear MMSE detector have been found. Based on this result, it is suggested that with sufficiently low system load 

\[ \beta \leq \frac{1}{L+1}, \]

where \( L \) is the number of effective interfering cells, a better performance than the fully loaded system may be attained in the strong interference environment with sufficiently large signal power. If the interference is weak, then a high-load system with fairly simple group linear MMSE detection has the advantage.

**Chapter 6**: The potential benefits of multiuser detection for crosstalk mitigation in DMT VDSL systems have been shown. ML-MUD can essentially eliminate crosstalk signals in DMT systems at a cost of high complexity. As a tradeoff, IC-MUD can significantly outperform single-user detection, with lower complexity than ML-MUD. The effectiveness of the M-estimator in combating the impulse noise has also been shown. Further, a new coded DMT VDSL receiver structure using turbo multiuser detection has been proposed and has been shown to achieve an overall 7.0 dB gain over the uncoded optimum (maximum likelihood) receiver at BER \( 10^{-7} \) for a channel with severe ISI, AWGN, and one dominant crosstalk signal. The detrimental effect of impulse noise can be overcome through an erasure decoding technique.

Several considerations for future work are now in order.

- As has been shown in Chapters 4 and 5, there is still a remarkable gap between the achievable capacity and the single cell upper bound with strong co-channel interference. Therefore, it is worth exploring the cooperation of base
stations and the feedback of channel information to enhance the system capacity in this case. Several promising techniques of interest are: joint transmission [6], ranked known interference [97], and joint transmit and receive optimization [91].

- The capacity of the whole downlink cellular system, when modeled as a broadcast/interference channel, has long been an open question. Work on multiuser information theory in the context of multicell cellular communication systems is also of interest.

- Besides the physical layer processing focused in this dissertation, the study of MAC and network layers, and opportunities for cooperation between these layers and the physical layer in the context of wireless LAN IEEE 802.11, is of interest. So is the potential for interdisciplinary collaboration in fields such as vehicle transportation, multimedia communications, and deep-space communications.
Bibliography


[Huaiyu Dai’s publications]


