Blind Multi-user Cancellation using the Constant Modulus Algorithm

A dissertation submitted in partial fulfilment of the requirements
for the degree

Master of Engineering (Electronic Engineering)

in the Department of Electrical, Electronic and Computer Engineering

in the

School of Engineering

at the

UNIVERSITY OF PRETORIA

Advisor: Professor L.P. Linde

Johan Pieter de Villiers

July 24, 2003

SUMMARY

Constant Modulus Detectors for Blind Multiuser Detection

by

J.P. de Villiers

Advisor: Professor L.P. Linde

Department of Electrical, Electronic and Computer Engineering

Master of Engineering (Electronic)

Keywords:

Blind Multiuser Detection, Linearly Constrained Constant Modulus Algorithm, Linearly Constrained Differential Constant Modulus Algorithm, Blind Equalization, DS-CDMA

Multiuser detection in direct sequence code division multiple access (DS-CDMA) systems has received much attention in recent years. This activity can be attributed to the fact that DS-CDMA systems are to be used in third generation (3G) cellular networks. Cellular operators have already paid billions of dollars for 3G licences, and are serious about ensuring effective use of available channel resources. It is for this reason that methods for increasing channel capacity are being vigorously researched. Multiuser detection can increase effective channel usage significantly, and so save the operators large amounts of money. The multiuser detection problem is discussed in the following paragraph.

In a DS-CDMA channel, each user is separated from the other users by his/her unique signature waveform. These signature waveforms are in practice quasi-orthogonal, i.e. exhibiting small amounts of cross correlation between the signature waveforms of different users. This has an adverse effect on performance, and introduces what is termed as *multiuser interference* when demodulated with a conventional matched filter detector. The effect of multiuser interference is especially visible in a channel consisting of unequal power users. Multiuser detection techniques concern themselves with the minimization of multiuser interference. Many techniques have been considered, of which the linear adaptive detectors offer good performance while employing relatively simple structures. A well known linear adaptive detector is the adaptive minimum mean square error (MMSE) detector.

This detector offers a significant improvement when compared with the conventional matched filter detector. Furthermore, it is adaptive, and is able to follow slow variations in the channel. The adaptive MMSE detector has the disadvantage that training sequences need to be transmitted to allow the detector to initially converge. In addition, if large channel fluctuations occur, the detector has to be retrained. These training sequences sacrifice valuable bandwidth, and are undesirable. This poses the need for blind detectors, that are able to adaptively tune out multiuser interference without the need for training sequences. The application of the widely used blind equalization constant modulus algorithm (CMA) to blind multiuser detection is attempted in this dissertation.

Direct application of the constant modulus algorithm to blind multiuser detection poses two distinct problems. The first is the fact that detecting any other user rather than the desired user may yield a constant modulus signal. This means that the standard constant modulus detector may lock onto any of the active users in the channel. The second problem is that some of the desired user's signal may be cancelled, even if the receiver locks onto the desired user component. Both of these problems may be solved by implementing a linear constraint which restricts the constant modulus detector to operate only on the subspace orthogonal to the desired user component. This detector is called the linearly constrained constant modulus (LCCM) detector. This detector exhibits performance equalling that of the MMSE detector subject to the fact that the desired user component is greater than a fixed value. This limitation may be a problem, especially in the case where the different users may have greatly varying amplitudes, such as in a mobile fading channel. The linearly constrained differential constant modulus (LCDCM) detector is the solution to this problem. The LCDCM detector penalizes any deviation in signal modulus from one sample to the next, whereas the LCCM detector penalizes any deviation in signal modulus from a constant value. The LCDCM detector has no limitation on minimum desired user amplitude, and convergence of the adaptation algorithm is assured.

In this dissertation, the two constant modulus multiuser detectors are analyzed, evaluated and compared with the MMSE detector. Existing signal and channel models are expanded to encompass the complex valued multipath DS-CDMA channels. For the first time, a global convexity condition is derived for the LCCM detector cost function. Simulation results for different channel types are generated and discussed. These channel types range from the additive white Gaussian noise (AWGN) channel to multipath fading channels.

OPSOMMING

Konstante Modulus Detektors vir Blinde Multigebruiker Deteksie

deur

J.P. de Villiers

Studieleier: Professor L.P. Linde

Departement Elektriese, Elektroniese en Rekenaar Ingenieurswese

Meester in Ingenieurswese (Elektronies)

Sleutelwoorde: Blinde Multigebruikerdeteksie, Lineêr Beperkde Konstante Modulus Algoritme, Lineêr Beperkde Differensiële Konstante Modulus Algoritme, Blinde Vereffening, DS-CDMA

Multigebruiker-deteksie in direkte sekwensie kodeverdeling multi-toegang (DS-KVMT) stelsels het baie aandag in die laaste paar jare ontvang. Hierdie aktiwiteit kan toegeskryf word aan die feit dat DS-KVMT stelsels in derde generasie (3G) sellulêre netwerke gebruik gaan word. Sellulêre operateurs het reeds biljoene dollar betaal vir 3G lisensies. Dit is om hierdie rede dat metodes om kanaalkapasiteit te vermeerder met so baie toewyding nagevors word. Multigebruiker-deteksie kan kanaalverbruik merkwaardig vermeerder, en kan dus die operateurs groot hoeveelhede geld bespaar. Die multigebruiker-deteksie-probleem word in die volgende paragraaf bespreek.

In 'n DS-KVMT kanaal word elke gebruiker van 'n ander een geskei d.m.v. sy/haar unieke identifikasiegolfvorm. Hierdie identifikasiegolfvorms is in die praktyk kwasi-ortogonaal, m.a.w. klein hoeveelhede kruiskorrelasie bestaan tussen die identifikasie golfvorms van die verskillende gebruikers. Hierdie eienskap het 'n nadelige effek op werkverrigting, en stel die kanaal bloot aan *multigebruiker-oorvleueling* wanneer dit d.m.v. 'n aangepaste filter gedemoduleer word. Die effek van multigebruiker oorvleueling is veral sigbaar in 'n kanaal wat uit gebruikers met ongelyke drywing bestaan. Multigebruiker deteksie tegnieke poog om die hoeveelheid multigebruiker oorvleueling binne 'n kanaal te minimeer. Baie tegnieke is al oorweeg, waarvan die lineêr aanpasbare detektors goeie werkverrigting lewer terwyl dit van 'n eenvoudige struktuur gebruik maak. 'n Bekende lineêr aanpasbare detektor is die minimum gemiddelde kwadraat fout (MGKF) detektor. Hierdie detektor bied 'n merkwaardige

verbetering wanneer vergelyk word met die konvensionele aangepaste filter detektor. Verder, is dit ook aanpasbaar, en het die vermoë om stadige kanaalveranderings te volg. Die aanpasbare MGKF detektor het die nadeel dat opleidingsekwensies nodig is om die detektor aanvanklik toe te laat om te konvergeer. Daarbenewens, as groot kanaalvariasies plaasvind, moet die detektor weer geleer word. Hierdie opleidingsekwensies offer duursame bandwydte op, en is dus ongewens. Dit stel die behoefte aan blinde detektors daar, wat aan kan pas om multigebruiker oorvleueling te minimeer sonder 'n behoefte aan opleidingsekwensies. Die toepassing van die algemeen gebruikte blinde vereffening konstante omhulling algoritme op blinde multigebruiker deteksie word in hierdie verhandeling aangespreek.

Die direkte toepassing van die konstante omhulling algoritme op blinde multiverbruiker deteksie bring twee spesifieke probleme mee. Die eerste is deteksie van enige ander gebruiker buiten die gewensde gebruiker sal ook 'n konstante omhulling lewer. Dit beteken dat die gewone konstante omhulling detektor kan sluit op enige van die aktiewe gebruikers in die kanaal. Die tweede probleem is dat 'n gedeelte van die gewensde gebruiker sein uitgekanselleer kan word, selfs al sluit die ontvanger op die gewensde gebruiker komponent. Beide hierdie probleme kan opgelos word deur gebruik te maak van 'n lineêre beperking, wat die werking van die konstante omhulling detektor beperk tot die subruimte ortogonaal tot die gewensde gebruiker komponent. Hierdie detektor word die lineêr beperkte konstante omhulling (LBKO) detektor genoem. Hierdie detektor lewer dieselfde werkverrigting as die MGKF detektor, onderhewig aan die beperking dat die gewensde gebruiker komponent groter as 'n spesifieke waarde is. Hierdie beperking mag 'n probleem wees, veral in gevalle waar die verskillende gebruikers grootliks varieërende amplitudes mag hê, soos in 'n mobiele deinende kanaal. Die lineêr beperkte differensiële konstante omhulling (LBDKO) detektor is die oplossing tot hierdie probleem. Die LBDKO detektor penaliseer enige afwyking in seinomhulling vanaf een monster tot die volgende, terwyl die LBKO detektor enige afwyking in omhulling vanaf 'n konstante waarde penaliseer. Dit het die gevolg dat die LBDKO detektor geen minimum beperking op gewensde geruiker amplitude het nie, en dat konvergensie van die aapassingsalgoritme verseker is.

In hierdie verhandeling word daar 'n analise en evaluasie van die twee konstante omhulling detektors gedoen, en word hulle met mekaar vergelyk, asook die MGKF detektor. Bestaande sein en kanaal modelle word uitgebrei om komplekse waarde multipad DS-KVMT kanale te akkomodeer. Vir die eerste keer word 'n globale konveksiteitsvoorwaarde vir die LBKO detektor kostefunksie afgelei. Simulasie resultate vir verskillende kanaaltipes word gegenereer en bespreek. Hierdie kanaaltipes wissel van die sommeerbare wit Gaussiese ruis (SWGR) kanaal tot multipad deinende kanale.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my Creator. I am in boundless debt to Him for my salvation, and all the wonderful talents I received.

I would like to thank my study leader Prof. Louis Linde who supportively guided me in the completion of my dissertation. Without his help and leadership, this project would not have been possible. Prof. Linde's dedication and self discipline contribute much to the field of digital communications, in the form of research outputs and industry projects. These qualities, together with his ability to envision novel opportunities and possibilities, make him a valuable asset to the department.
Further thanks must go my parents, whom in their love, encouraged me to continuously educate myself. Without their financial and moral support, I would not have been able to succeed up to this level.

My wife, Helena provided with much support while I was working on this dissertation. For this I am very grateful.

For further financial support, through a fellowship, I would like to extend my thanks to Research Enterprises at the University of Pretoria.

Special thanks must go to Professors E.E. Rosinger and J.A. Snyman of the University of Pretoria mathematics and mechanical engineering departments, who patiently helped me with aspects of optimization theory. The following people also helped me in some other aspects of this dissertation. These, in no particular order, include: Whasuck Lee, Conrad Beyers, Jacques Cilliers, Jacques van Wyk, Francois Pienaar and Leon Staphorst.

TABLE OF CONTENTS

CHAPTER ONE - INTRODUCTION 1
1.1 Introduction . . . 1
1.1.1 General Problem Definition . . . 3
1.2 An Overview of Code Division Multiple Access (CDMA) . . . 4
1.3 The CDMA Data Detection Hierarchy . . . 4
1.3.1 Single User Detection . . . 5
1.3.2 Multiuser Detection . . . 5
1.3.2.1 Joint Detection (JD) . . . 7
1.4 Applying the Constant Modulus Criterion to MUD . . . 8
1.5 Contributions of this Dissertation . . . 9
1.6 Outline of Dissertation Chapters . . . 11

CHAPTER TWO - SYNCHRONOUS AND ASYNCHRONOUS CDMA MODELS 13
2.1 The CDMA Signal Model . . . 13
2.2 Discrete-time Synchronous Models . . . 15
2.2.1 Matched Filter Outputs . . . 15
2.2.2 Whitened Matched Filter Model . . . 17
2.2.3 Orthonormal Projections . . . 18
2.3 Discrete-time Asynchronous Models . . . 19
2.3.1 Intersymbol Interference . . . 20
2.3.2 Asynchronous Vector Matrix Model . . . 21
2.4 The Fading Mobile Channel Model . . . 24
2.4.1 Rayleigh Fading due to Doppler Spread - Clarke's Model . . . 25
2.4.2 Multipath Time Dispersion Model . . . 25
2.4.2.1 Discrete Time Channel Impulse Response . . . 26
2.4.2.2 The CDMA Uplink and Downlink Channels . . . 28
2.5 Summary . . . 28

CHAPTER THREE - THE MATCHED FILTER RECEIVER AND MULTIUSER DETECTION PERFORMANCE MEASURES 29
3.1 Optimal Decision Rules and Sufficient Statistic . . . 29
3.1.1 Decision Rules and Decision Regions . . . 29
3.1.2 Continuous-Time Gaussian Signals . . . 31
3.1.3 Sufficient Statistic . . . 32
3.2 The Optimal Receiver - Single User . . . 33
3.2.1 Linear Detectors . . . 33
3.2.2 Error Probability - Optimal Single User Linear Detector . . . 34
3.2.3 Error Probability - Optimal Single User Non-Linear Detector . . . 36
3.3 Matched Filter Error Probability - Synchronous Users . . . 37
3.3.1 The Two User Case . . . 38
3.3.1.1 BEP as performance measure - The two user case . . . 41
3.3.1.2 The two user signal space representation . . . 43
3.3.2 The K-user Case . . . 46
3.3.3 The Gaussian Approximation for BEP . . . 49
3.4 Matched Filter Error Probability - Asynchronous Users . . . 53
3.5 Asymptotic Multiuser Efficiency and Related Measures . . . 54
3.5.1 Asymptotic Multiuser Efficiency of the Two User Matched Filter . . . 57
3.5.2 Asymptotic Multiuser Efficiency of the K User Matched Filter . . . 58
3.6 Performance of the Coherent Single User Matched Filter Detector in Frequency Flat Fading 60
3.6.1 The Single User Case in the Presence of Fading . . . 60
3.7 Summary . . . 62

CHAPTER FOUR - LINEAR MULTIUSER DETECTORS 64
4.1 The Linear Decorrelating detector . . . 64
4.2 The Optimum Linear Detector . . . 65
4.2.1 The Two User Optimum Linear Detector . . . 66
4.3 The Linear MMSE Detector . . . 68
4.3.1 The MMSE Optimization Problem . . . 69
4.3.2 The MMSE Detector Vector Matrix Model . . . 69
4.3.3 The Two User MMSE Detector . . . 72
4.3.4 The Limiting Forms of the MMSE Detector . . . 72
4.3.5 The Asynchronous MMSE Detector . . . 73
4.3.6 The Wiener Filter Characterization of the MMSE Detector . . . 73
4.4 The MMSE Detector Least Mean Square (LMS) Algorithm . . . 75
4.5 Performance of the MMSE Detector . . . 78

4.5.1 Signal-to-Interference Ratio of the MMSE Detector . . . 78
4.5.2 Asymptotic Multiuser Efficiency and Near-Far Resistance of the MMSE Detector . 79
4.5.3 BEP of the MMSE Detector . . . 80
4.5.3.1 Gaussian Approximation of the MMSE Detector BEP . . . 82
4.5.3.2 Infinite User Limit of the MMSE Detector BEP . . . 83
4.5.4 Power Tradeoff Regions of the MMSE Detector . . . 84
4.5.5 MMSE Detector Performance in Multipath Channels . . . 84
4.5.5.1 Signal-to-interference Ratio of the MMSE detector in a Multipath Channel 90
4.5.5.2 BEP of the MMSE Detector in a Multipath Channel . . . 91
4.6 Summary . . . 91

CHAPTER FIVE - BLIND MULTIUSER DETECTION USING THE CONSTANT MODULUS CRITERION 93
5.1 Introduction . . . 93
5.2 The Linearly Constrained Constant Modulus Criterion . . . 95
5.2.1 The Convexity of the LCCM Cost Function . . . 96
5.2.2 The Stationary Points of the LCCM Cost Function . . . 98
5.2.3 Linearly Constrained Constant Modulus Algorithm . . . 102
5.3 The Linearly Constrained Differential Constant Modulus Criterion . . . 106
5.3.1 Global Minimum of the LCDCM Cost Function . . . 108
5.3.2 Linearly Constrained Differential Constant Modulus Algorithm . . . 109
5.4 Performance of the LCCM and LCDCM Algorithms in Multipath Fading Channels . . . 111
5.5 Summary . . . 112

CHAPTER SIX - SIMULATION RESULTS OF THE CM DETECTOR 114
6.1 Simulation Setup . . . 114
6.1.1 Single Path AWGN Channel . . . 116
6.1.2 Static Multipath Channel . . . 116
6.1.3 Single- and Multipath Fading Channels . . . 116
6.2 Performance in an AWGN Channel . . . 117
6.3 Performance in a Static Multipath Channel . . . 123
6.4 Performance in Rayleigh Fading Single- and Multipath Channels . . . 126

CHAPTER SEVEN - FINAL SUMMARY AND CONCLUSIONS 130
7.1 Theoretical Summary and Conclusions . . . 130
7.2 Simulation Summary and Conclusions . . . 132

7.3 Proposals for Further Research 133

REFERENCES 134

APPENDIX A - SELECTED PROPERTIES OF THE Q-FUNCTION 140

APPENDIX B - SIMULATION OF MOBILE CHANNEL 142
B.1 Doppler Spread Rayleigh Fading 142
B.2 Frequency Selective Multipath Rayleigh Fading 143

APPENDIX C - DIFFERENTIATION WITH RESPECT TO A COMPLEX MATRIX 146
C.1 Basic Definitions 146
C.2 The Gradient Matrix in Terms of the Derivative with Respect to a Matrix 148
C.3 Differentiating the Components of the MMSE Cost Function 149

APPENDIX D - EVALUATING THE EXPECTED VALUE IN THE LCCM AND LCDCM COST FUNCTIONS 152
D.1 Expected Value in the LCCM Cost Function 152
D.2 Expected Value in the LCDCM Cost Function 153

LIST OF FIGURES

1.1 Data detection hierarchy structure for CDMA 6
1.2 Graphical representation of the structure and outline of the dissertation. 12
2.1 Schematic representation of the cross correlation between two synchronous users 14
2.2 Block diagram illustration of the complex matched filter receiver 16
2.3 Block diagram description of the orthonormal projection correlation receiver 18
2.4 Schematic representation of the symbol epochs for three users if $M = 1$ 20
2.5 Block diagram of the z domain vector matrix model of equations (2.36) and (2.42) . . . 23
2.6 The mobile radio channel as a function of time and space. 25
2.7 The time varying discrete time impulse response model for a specific multipath radio channel. 27
3.1 Conditional distributions of Y given $b = -1$ and $b = +1$ 35
3.2 Block diagram depicting the bank of matched filters for multiple CDMA users 37
3.3 Block diagram depicting the special case of the two user CDMA matched filter receiver structure 38
3.4 Output of the matched filter with one interfering user and $A_2/A_1 > 1/|\rho|$ 40
3.5 BEP of the matched filter detector for different relative amplitudes and $\rho = 0.2$ 41
3.6 Regions of signal-to-noise ratios to attain a BEP of 3×10^{-5} for both users 42
3.7 Signal space diagram in the (y_1, y_2) space for equal amplitudes and $\rho = 0.2$ 44
3.8 Joint probability density function in the (y_1, y_2) space for equal amplitudes, $\rho = 0.2$ and $\sigma = 1$ 45
3.9 Overhead view of the joint probability density function in the (y_1, y_2) space for equal amplitudes, $\rho = 0.2$ and $\sigma = 1$ 46
3.10 Signal space diagram in the alternative orthogonal $(\tilde{y}_1, \tilde{y}_2)$ space for equal amplitudes $A_1 = A_2$ and $\rho = 0.2$ 47
3.11 Signal space diagram in the alternative orthogonal $(\tilde{y}_1, \tilde{y}_2)$ space for $A_2 = 6A_1$ and $\rho = 0.2$ 48
3.12 BEP as a function of SNR with $K = 10$ equal energy users and $\rho = 0.08$ (eye open) . . 50
3.13 BEP as a function of SNR with $K = 14$ equal energy users and $\rho = 0.08$ (eye closed) . . 51
3.14 Asymptotic multiuser efficiency for a matched filter detector with two equal energy users and $\rho = 0.2$ 59

4.1 Block diagram depicting the structure of the K user linear receiver performing a linear operation $\mathbf{M}$ on the sampled matched filter outputs. 65
4.2 Block diagram depicting the structure of the two user linear decorrelating, optimum and MMSE receivers in the real domain. 67
4.3 Asymptotic Multiuser Efficiencies of the Matched Filter, Decorrelating and MMSE Detectors. 81
4.4 BEP graph comparing the exact and Gaussian approximated curves of the MF and MMSE detectors. 83
4.5 Regions of signal-to-noise ratios to attain a BEP of 3×10^{-5} for both users using a MMSE detector. 85
4.6 Depiction of the equivalent synchronous multipath model of a CDMA channel. 86
5.1 Complex LCCM cost function surface with $A_1^*A_1 > \alpha/4$. 102
5.2 Complex LCCM cost function surface with $A_1^*A_1 = \alpha/4$. 103
5.3 Complex LCCM cost function surface with $A_1^*A_1 < \alpha/4$. 104
5.4 Generalized Sidelobe Canceller with $\mathbf{x}[i]$ governed by (5.53) in the case of the LCCM algorithm and (5.79) in the case of the LCDCM algorithm. In the case of the LCDCM algorithm, the previous values $Z_D[i]$, $Z_{MF_D}[i]$ and $\mathbf{r}_D[i]$ need to be remembered to compute $\mathbf{x}[i]$. 105
5.5 One Dimensional Complex LCDCM cost function surface. 110
6.1 System block diagram of the simulation setup. 115
6.2 Signal to noise and interference ratios versus time of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the MMSE, LCCMA and LCDCMA detection techniques. 118
6.3 Signal to noise and interference ratios versus time of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the MMSE, LCCMA and LCDCMA detection techniques. In this case $A_1^*A_1 < \alpha/4$. 119
6.4 Signal to noise and interference ratios versus time of a CDMA system with 3 users, and a 34dB strong fourth user powering on at time $t = 4000$ symbols. 120
6.5 Signal to noise and interference ratios versus no. of users of a CDMA system with a spreading factor of 7 and $E_b/N_0 = 10$ dB in an AWGN channel. 121
6.6 Bit error rate versus no. of users of a CDMA system with a spreading factor of 7 and $E_b/N_0 = 10$ dB in an AWGN channel. 122
6.7 Signal to noise and interference ratios versus time of a CDMA system with K = 6 users and a SF = 7 in an AWGN channel. The plot shows the performance for a code mismatch with mismatch variance of 0.1. 123
6.8 BER of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques. 124

6.9 Z-plane plot of a minimum phase static 3-ray multipath channel. The multipath profile consists of 0.86 at zero delay, 0.43 at 1 chip delay and 0.26 at 2 chips delay. 125
6.10 Frequency response of the static three-ray multipath channel of which the z-plane representation is shown in Figure 6.9. 126
6.11 BER of a CDMA system with 6 users and a spreading factor of 7 in a static 3-ray multipath channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques. The multipath profile consists of 0.86 at zero delay, 0.43 at 1 chip delay and 0.26 at 2 chips delay. 127
6.12 BER of a CDMA system with 6 users and a spreading factor of 7 in a Rayleigh fading single-path channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques. 128
6.13 BER of a CDMA system with 6 users and a spreading factor of 7 in a 3-ray Rayleigh fading multipath channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques. 129
7.1 Knowledge needed for the different types of multiuser detection schemes. 131
B.2 Baseband complex Rayleigh fading coefficient simulator. 142
B.3 Frequency spectrum of 3rd order approximation of a Doppler filter with a Doppler frequency of 50Hz. 144
B.4 Model of a frequency selective (multipath) fading channel. 145

List of Abbreviations and Acronyms

A

AWGN - Additive White Gaussian Noise

B

BEP - Bit Error Probability
BER - Bit Error Rate
BPSK - Binary Phase Shift Keying

C

CDMA - Code Division Multiple Access
CMA - Constant Modulus Algorithm
CM - Constant Modulus

D

DD - Decision Directed
DF - Decision Feedback
DFE - Decision Feedback Equalizer
DS - Direct Sequence
DSP - Digital Signal Processor

E

EVD - Eigenvalue Decomposition

F

FIR - Finite Impulse Response
FPGA - Field Programmable Gate Array

G

GCL - General Chirp Like (sequences)

I

IC - Interference Cancellation
ISI - Inter Symbol Interference

J

JD - Joint Detection

L

LCCM - Linearly Constrained Constant Modulus
LCCMA - Linearly Constrained Constant Modulus Algorithm
LCDCM - Linearly Constrained Differential Constant Modulus
LCDCMA - Linearly Constrained Differential Constant Modulus Algorithm
LCMV - Linearly Constrained Minimum Variance
LCMVA - Linearly Constrained Minimum Variance Algorithm
LMS - Least Mean Square

M

MAI - Multiple Access Intereference
MF - Matched Filter
MIMO - Multiple Input Multiple Output
MISO - Multiple Input Single Output
ML - Maximum Likelihood
MLSE - Maximum Likelihood Sequence Estimation
MMSE - Minimum Mean Square Error
MOE - Minimum Output Energy
MSE - Mean Square Error
MUD - Multi User Detection

N

MAI - Narrow Band Interference

R

RU - Root of Unity

S

SD - Sequence Detection
SNR - Signal to Noise Ratio
SSD - Single Symbol Detection
SISO - Single Input Single Output

SIR - Signal to (Noise and) Interference Ratio
SUD - Single User Detection
SVD - Singular Value Decomposition

T

TDMA - Time Division Multiple Access

V

VA - Viterbi Algorithm

Z

ZF - Zero Forcing

LIST OF SYMBOLS

$\mathbf{A}$ - Matrix of which the diagonal contains all the users' amplitudes

A - The single user channel amplitude

$a_k(n)$ - Pseudo noise sequence of the signature waveform of user k

A_k - The received signal amplitude of user k

$\tilde{A}_k$ - The complex valued amplitude of user k due to a phase θ_k

$A_{k,l}$ - The received signal amplitude of the lth multipath component of user k

$\bar{A}^2$ - The statistical average of the squared amplitudes of all interfering users

$\mathbf{b}$ - Vector of all K users' transmitted bits

b_k - The bit transmitted by the kth user

$\hat{b}$ - The bit decision of a single user channel

$\hat{b}_k$ - The bit decision of the kth user

$b_k[i]$ - The ith bit transmitted by the kth asynchronous user

$\mathbf{c}$ - An arbitrary K-vector $\mathbf{c} = (c_1, \ldots, c_K)^T$.

c_k - Complex waveform of duration T used for a linear transform

c_k^s - Subspace of c_k spanned by the signature waveforms $s_1, \ldots, s_K$

c_k^o - Subspace of c_k spanned orthogonal to the signature waveforms $s_1, \ldots, s_K$

$\mathbf{C}$ - Covariance matrix of the received signal vector $\mathbf{r}$

f_{d_k} - The Doppler frequency of user k

$e_k(\sigma)$ - Effective energy of user k

E_b - Energy per bit

$\mathbf{F}^{-H}$ - Whitening filter

G - The subset of interferers with the partially open eye condition satisfied

$\bar{G}$ - The subset of interferers with the closed eye condition satisfied

$h_k(t)$ - Linear time invariant channel impulse response of user k

$h_k(t,\tau)$ - Linear channel impulse response of user k

H_i - The ith hypothesis out of a total of m hypotheses

$\mathbf{H}(J)$ - The Hessian matrix of the cost function J

$\mathbf{I}$ - Identity matrix

J - Joint cost function of all K users

J_k - Cost function of user k

$J_{\min}$ - Minimum mean square error of the Wiener filter

K - The number of simultaneous users

L - The number of dimensions or orthonormal signals in the orthonormal detector model

M - Asynchronous user packet length is equal to $(2M+1)$

$\mathbf{M}$ - Arbitrary linear transform on the received signal vector

$\bar{\mathbf{M}}$ - Optimum linear transformation in a minimum mean square error sense

$\tilde{\mathbf{M}}$ - Scaled version of an optimum linear transformation in a MMSE sense

$\tilde{\mathbf{M}}_a$ - Asynchronous channel optimum linear transformation in a MMSE sense

$\bar{\mathbf{n}}$ - Vector of K whitened (uncorrelated) noise components

$\mathbf{n}$ - The vector which contains the noise components of all K matched filter outputs

n_k - The unnormalized inner product between $n(t)$ and the signature waveform of user k

$n_k[i]$ - The ith bit asynchronous noise contribution of user k

$n(t)$ - White Gaussian noise with unit power spectral density

N - The length (or spreading gain) of the pseudo noise sequence for each signature waveform

N_0 - The one sided noise spectral density

$\mathcal{N}(\mu,\sigma^2)$- Notation for a Gaussian distribution with mean μ and variance σ^2

$\mathbf{p}$ - The cross correlation vector between the vector $\mathbf{r}$ and the desired response b_1

$p(t)$ - The chip waveform of duration T_c

P - Number of equally spaced discrete multipath components

$\mathcal{P}$ - Maximum allowable bit error probability in power-tradeoff region

P_e - Single user error probability in a Gaussian channel

$P_{e,R}$ - Error probability in a Rayleigh fading channel

$P_e(\sigma, k)$ - Error probability in a Gaussian channel for user k

$\tilde{P}_e(\sigma, k)$ - Gaussian approximation of error probability in a Gaussian channel for user k

$P_e^F(\sigma, k)$ - Error probability in a Rayleigh fading channel for user k

$P_e^{\mathbf{t}_k}$ - Error probability of user k due to a linear transform $\mathbf{t}_k$

$Q(\cdot)$ - The Q-function as defined in Appendix A

$R_p(\tau)$ - Autocorrelation function of an arbitrary rectangular chip signature waveform

$\mathbf{r}$ - The output vector of a correlation receiver

$\mathbf{R}$ - Synchronous cross correlation matrix

$\mathbf{R}[0]$ - Asynchronous cross correlation matrix as defined in (2.38)

$\mathbf{R}[1]$ - Asynchronous cross correlation matrix as defined in (2.39)

$\mathcal{R}$ - An arbitrary interval on the real line

R_i - The ith decision region out of a total of m regions

$s(t)$ - Unit energy deterministic signature waveform in a single user channel

$s_k(t)$ - Unit energy deterministic signature waveform of user k

$\tilde{s}_k(t)$ - Linear time dispersive channel signature waveform response of user k

$\mathbf{s}_k$ - Signature vector of user k

$\mathbf{S}$ - Signature matrix containing the signature vectors of K users

$S(z)$ - z-transform transfer function of the asynchronous cross correlations

T - Symbol duration equals $L \times T_c$

T_c - Chip duration

$\mathbf{v}$ - Arbitrary tap weight vector of a linear detector

$\bar{\mathbf{v}}$ - MMSE optimum tap weight vector

$x_i(t)$ - An arbitrary deterministic energy function defined on an interval $\mathcal{R}$ on the real line

$\mathbf{y}$ - Vector of K matched filter outputs

$\bar{\mathbf{y}}$ - Vector of K whitened matched filter outputs

y_k - Matched filter output of user k

$y_k[i]$ - Matched filter output of the ith bit of asynchronous user k

$y(t)$ - Received signal embedded in AWGN

Y - Decision statistic of $y(t)$

$\nabla_{m_{vw}}$ - Element of the vth row and wth column of the complex gradient operator matrix $\nabla_{\mathbf{M}}$

$\nabla_{\mathbf{M}}(J)$ - Complex gradient matrix of the cost function J

α - Arbitrary real scalar used as desired modulus in the constant modulus type detector

$\phi_{k,l}(t,\tau)$ - Phase shift of the lth multipath component due to fading and other channel effects

β - Ratio of number of users to spreading gain (K/N)

φ - A deterministic energy signal of duration T

γ - Signal-to-noise ratio (SNR)

γ_{ck} - Signal-to-interference ratio (SIR) of user k

η_k - Asymptotic multiuser efficiency of user k

$\hat{\eta}_k$ - Multiuser efficiency of user k

$\bar{\eta}_k$ - Near-far resistance of user k

η_k^F - Asymptotic multiuser efficiency of user k in a Rayleigh fading channel

$\mu_{\max}$ - Maximum step size to ensure convergence of the LMS algorithm

$\theta_{k,l}$ - Combined phase term of the lth multipath component of user k

Θ - The parameter to be inferred in a statistical inference problem

ρ_{kj} - Synchronous cross correlation between the signature waveforms of user k and user j

$\rho_{kj}(\tau)$ - Asynchronous cross correlation between two signature waveforms as defined in (2.7)

$\rho_{jk}(\tau)$ - Asynchronous cross correlation between two signature waveforms as defined in (2.8)

σ^2 - Additive white Gaussian noise variance

$\Delta\tau$ - The excess delay bin spacing

τ_l - The relative delay of the lth multipath bin of any user

τ_k - The time delay of asynchronous user k

ω - Worst asymptotic multiuser efficiency among all users

$\mathbf{\Omega}$ - Covariance matrix of the interference

$\tilde{\mathbf{\Omega}}$ - Covariance matrix of the interference in the multipath case

ψ_l - The lth orthonormal signal out of a set of L

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

The market of wireless communications is continuously experiencing rapid growth. Companies have already spent an enormous amount of money on third generation mobile licences. The leap of faith by such companies indicate a belief that there is considerable growth potential in the mobile communications sector.

The proposed third generation systems must deliver high speed data and voice services, while remaining compatible with second generation systems. Furthermore, new techniques of modulation, coding, equalization, multipath combining, multiuser detection, antenna and spatial diversity, and other mobile radio techniques are required to broaden the variety of existing services. There were several proposals for the new third generation cellular standard, of which the Wide band Code Division Multiple Access (W-CDMA) systems were taken the most seriously. The reason for this, is that W-CDMA offers increased flexibility when compared with the Time Division Multiple Access (TDMA) schemes that are predominant in current second generation cellular systems.

These recent developments in the cellular telecommunications market stimulated much research on how to increase system capacity in CDMA systems. One way in which the system capacity can be increased within a CDMA system is by minimizing the interference caused by other users. This type of interference, called Multiple Access Interference (MAI), can be limited by utilizing different multiuser detection techniques. In addition, frequency selectivity in the channel introduces another type of interference called, Inter Symbol Interference (ISI).

The field of MAI cancellation, which is collectively called multiuser Detection (MUD) [1], is rather

broad. The optimum multiuser detector derived by Verdu [2] attains single-user performance when the following is known:

1. The signature waveform of the desired user.
2. The signature waveform of the interfering user.
3. The timing of the desired user.
4. The timing of each of the interfering users.
5. The channel impulse responses of the desired user.
6. The channel impulse responses of the interfering users.

The optimum multiuser detector uses the maximum likelihood Viterbi Algorithm (VA) to do Maximum Likelihood Sequence Estimation (MLSE). There are distinct disadvantages when it comes to practical implementation of the optimum multiuser detector. Even for a small number of users, the computational complexity is enormous. A suboptimum approach is consequently needed.

Among the several different MUD methods there exists a class of suboptimum multiuser detectors which have an adaptive equalizer type structure (either linear or non-linear). The equalizer type detectors need to adapt to time varying radio channels, and use different criteria to do so. The most common of these are the decorrelating or zero forcing (ZF) and mean square error (MSE) criteria. The decorrelating or ZF criterion can be considered an asymptotic form of the MSE criterion. To initially adapt to the impulse response of the channel, training sequences are used. This means that only the desired user signature (1.), timing (3.) and impulse response (5.) in the above list needs to be known. When using training sequences, the channel impulse response can be estimated. Known training symbols are transmitted until accurate decisions can be made. After the training phase, the symbol decisions can be used to adjust the equalizer coefficients.

In [3], Honig proposed the use of a blind algorithm, based on the Minimum Output Energy (MOE) criterion, which eliminated the need for training sequences. There also exists a variety of other blind adaptation algorithms which is readily used for channel equalization, of which the Constant Modulus Algorithm (CMA) is the most widely used blind algorithm [4]. This dissertation will investigate the application of the Constant Modulus (CM) criterion to implement a blind equalizer multiuser detector. In this way, effective MAI cancellation can be achieved without the use of training sequences. This method will be implementable in CDMA systems that use constant modulus complex spreading sequences. An example of such sequences are the root of unity (RU) filtered generalized chirp like (GCL) sequences. [5, 6] Surprisingly, it has been shown that the constant modulus algorithm can also

be applied to non-constant envelope signals [4]. Applying the CMA to CDMA multiuser detection poses some new problems. To keep the CMA detector from capturing one of the interfering signals, a linear constraint must be imposed on the CMA cost criterion. The modified constant modulus criterion is termed the linearly constrained constant modulus algorithm (LCCMA).

In this dissertation the suitability of the LCCMA is investigated for the purpose of blind CDMA multiuser detection. The shortcomings of the LCCMA are identified, and improvements will be proposed.

1.1.1 GENERAL PROBLEM DEFINITION

The problem addressed in this dissertation can be defined as the investigation of the constant modulus criterion within the framework of multiuser detection in a linear decision directed (DD) equalizer structure, to achieve effective MAI cancellation in the uplink (asynchronous) and downlink (synchronous) of Direct Sequence (DS) CDMA systems.

Concerning system evaluation, the CDMA multiuser system must be evaluated within semi-static and mobile channel conditions by means of computer simulation. This is to be done by means of Bit Error Rate (BER) comparative testing as a function of bit energy to noise spectral density ratio E_b/N_0. In this way, the performance of the constant modulus technique can be compared with the matched filter detection bound (which will be explained later), as well as the multiuser channel single detection case. Furthermore, comparison with the standard Minimum Mean Square Error (MMSE) multiuser detection scheme will give a comparative measure of performance.

In summarizing, the dissertation objectives are

- The investigation of the constant modulus technique for effective MUD.
- The comparative C++ software simulation of the applicable constant modulus techniques in a single path static and a multipath mobile channel.
- Convergence performance by plotting signal to interference ratios versus time for different types of channels.
- BER performance evaluation of chosen techniques compared with the single user and MMSE multiuser detection cases in theory and software simulation.
- Investigation of other performance criteria such as asymptotic multiuser efficiency, signal to interference ratio, etc.

1.2 AN OVERVIEW OF CODE DIVISION MULTIPLE ACCESS (CDMA)

There are three CDMA categories that can be distinguished. These are direct sequencing (DS), time hopping (TH) and frequency hopping (FH). The FH and TH categories are generalizations of the FDMA (Frequency Division Multiple Access) and TDMA techniques respectively, in which the assignment of frequency bands and time slots are changed according to certain hop patterns. We will be mainly concerned with the DS category, which is based on spread spectrum principles.

The DS-CDMA system is a multiple access system in which many users are simultaneously multiplexed on the same frequency band by means of quasi-orthogonal codes. The data from each user is modulated by a technique called direct sequence spread spectrum (DSSS) modulation. In this method the modulated signal is generated by mixing a high frequency (chip rate) code sequence with the data at a much lower data rate. Some of the advantages of DS-CDMA, is the manner in which it can effectively reuse the available frequency resources and the efficient multiple access system capacity. Furthermore, the spread spectrum modulation scheme has an inherent immunity against multi-path mobile channels. This makes it ideal for cellular wireless mobile and semi-static channels.

By a proper choice of spreading codes, the cross correlation between the different users in a CDMA system can be minimized, thus decreasing the interference between different users. The residual interference can then be removed by means of multiple access interference cancellation techniques.

The basic DS-CDMA principle is based on conditions such as an additive white Gaussian noise (AWGN) channel with perfect power control. Acceptable power control is difficult to achieve in a time variant mobile fading channel. Imperfect power control causes the situation where the signal of a strong mobile completely overpowers weaker mobile signals. This situation is called the near-far effect, and is very detrimental to system capacity. It is thus important that multiuser detection techniques should be near-far resistant for mobile channels.

1.3 THE CDMA DATA DETECTION HIERARCHY

In order to form a proper heuristic view of all existing data detection principles, the references of Klein [1] along with Woodward and Vucetic [7] are of great value. The references [8] and [9] are successful in presenting a less detailed overall view of the CDMA multiuser detection problem. Because of the unified manner in which the set of multiuser detection algorithms are presented in [1], the hierarchical structure adopted in this reference will, with minor modifications, be presented in this section. Using this section, it is possible to see where the CM MUD method fits within the global CDMA data

detection hierarchy.

1.3.1 SINGLE USER DETECTION

The traditional signal separation by means of matched filters (MFs) is termed single detection in [1]. We will use the term single user detection (SUD) as to avoid confusion with the term sequence detection (SD). The SUD method assumes perfect power control of all users. The SUD method is suboptimal in that it treats all interference (both known and unknown) as noise. We know that this is not the case with CDMA in a mobile channel, as both MAI and inter-symbol interference (ISI) are not noise like. MAI is caused by cross correlation terms between the different users' spreading codes, while ISI is caused by the interference of several unequally delayed incident waves due to scatterers surrounding the receiver. There is a strong connection between MAI and ISI, and is explained in the reference [2].

The SUD method is in a practical sense not near-far resistant, as it assumes perfect power control. As mentioned earlier, this is very difficult to achieve in a time variant mobile fading channel.

1.3.2 MULTIUSER DETECTION

The poor system capacity in traditional matched filter signal separation (SUD), is the result of cross-correlation between user codes in synchronous systems, and the loss of orthogonality due to phase offsets in asynchronous systems. Recent advances in multiuser theory shows that the high MAI inherent in CDMA systems is not necessarily a shortcoming of the CDMA multiple access scheme. It is possible to use some or all of the code and channel information to cancel out MAI. This method of detection, of which many variants have already been proposed, is collectively called multiuser detection (MUD). We can subdivide MUD into two categories which are called interference cancellation (IC) and joint detection (JD). The hierarchical structure for data detection principles in CDMA is depicted in Figure 1.1.

INTERFERENCE CANCELLATION (IC)

The idea of IC is closely related to decision feedback (DF) and is

- to detect part of the transmitted data symbols,
- to reconstruct the contribution of these transmitted data to the compound received signal and
- to subtract the contribution from the compound received signal.

This means that there remains a component of MAI that is still treated as noise, thus making the joint detection principle suboptimum. There are currently two methods of interference cancellation. The methods of

- successive or serial IC and
- parallel IC

can be distinguished.

Successive IC sorts users from the strongest to the weakest signal and then detects the data symbol of the strongest user to cancel its influence. With the contribution of the strongest user cancelled out, the method then cancels the influence of the second strongest user. This is repeated until all of the users' influence is cancelled out. This method is ideally suited for the case of users with varying signal strengths [1, 9], i.e. the near-far effect. It is related to the decision feedback algorithm and is non-linear. Standard Parallel IC on the other hand, detects the contributions of all the users' data simultaneously at the output of each single user detector. The influence of all of the users, except the user in question, is cancelled from the received signal. In general, the parallel IC method has a higher potential of performance enhancement than successive IC, since the contributions of all the users are cancelled, and not only those which have stronger signal power. The parallel IC method can be repeatedly performed, leading to a *multistage canceller* default.

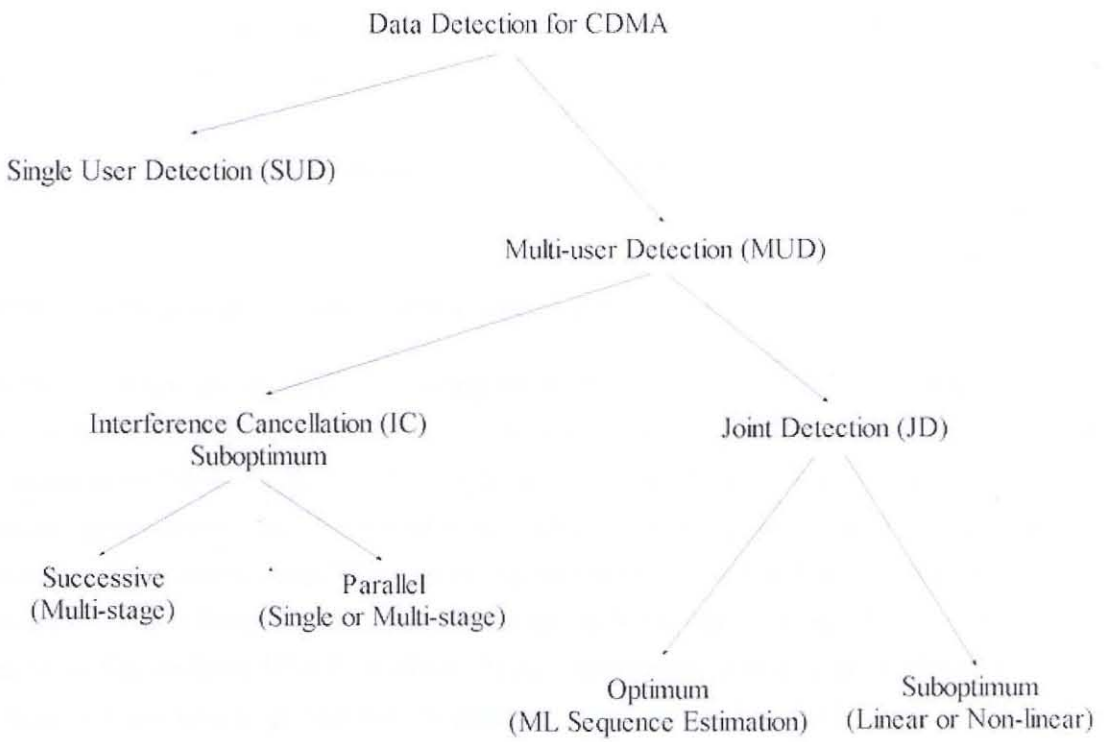

Figure 1.1: Data detection hierarchy structure for CDMA

1.3.2.1 JOINT DETECTION (JD)

The joint detection idea is based on the idea that the data symbols of all the users are detected jointly, using all the a priori knowledge about the MAI. We can divide joint detection into two groups. These are optimum and suboptimum detection. Both the optimum and suboptimum joint detection algorithms have a greater potential to enhance system capacity than IC methods.

1.3.2.1.1 Optimum Maximum Likelihood (ML) JD

The optimum detector for the multiuser CDMA system was first demonstrated by Verdu [2]. It was shown that the output of a bank of matched filters (MF), i.e. SUD, contains sufficient statistics to achieve optimum detection. This can be done by a Viterbi algorithm for maximum likelihood sequence estimation (MLSE). The practical application of this approach is limited by two main factors. The first is that the number of states required by the Viterbi algorithm is exponentially complex in the number of active users. For many users, the problem becomes intractable. Secondly, the MF filter bank is formulated in an AWGN channel. In a practical time varying mobile channel with ISI, the MF bank will have to be synthesized with knowledge of the activity, time and phase synchronization, spreading sequence, power, and channel conditions for each user. Much of this information is also required by the Viterbi algorithm.

There are numerous simplified algorithms that can replace the Viterbi algorithm, as proposed in [10–12]

1.3.2.1.2 Suboptimum (Linear and Non-linear) JD

This family of JD can be classified as being adaptive linear (Decision Directed - DF) or non-linear (decision feedback - DF) [1, 7, 9]. The linear type JD receivers perform a linear transformation on the output of the MF bank. The non-linear type receiver, has a forward filter operating directly on the received signal samples, and a backward (feedback) filter operating successively on a non-linear decision to cancel out interference. There are two minimization operations that can be implemented with both linear and non-linear JD techniques. These are the decorrelating or Zero-Forcing (ZF) and Minimum Mean Square Error(MMSE) methods. These combinations give us a group of four suboptimum JD receiver types which are near-far resistant. As mentioned earlier, the ZF and MMSE multiuser detectors have the disadvantage of needing to be trained either at the beginning of reception, or at regular intervals between blocks of data. Honig [3] was the first person to suggest blind suboptimum JD.

Blind Methods

Li [13], gives a summary of recent work on blind MUD. Honig [3] first proposed blind MUD using the minimum output energy criterion. Recently, some other methods for joint MUD and blind equalization were presented [13–23].

The first kind is subspace-based methods [14–16]. These methods usually require Singular Value Decomposition (SVD) or Eigenvalue Decomposition (EVD) of some data correlation matrix. These computations are complex and not very practical to implement. Another drawback of the subspace type methods, is that accurate rank determination is difficult in a noisy environment.

The second kind of blind MUD methods is constrained optimization [3,17,18], which result in computationally efficient adaptive algorithms. The methods in [3] and [18], are based on the MOE criterion, where as the CMA MUD is based on the constant modulus criterion. The major drawback of the MOE methods, is that there exists a saturation effect in the steady state, which causes a significant performance gap between the blind MOE detector and the true MMSE detector [3, 16]. Furthermore, the performance of the MOE method critically depends on the nonzero magnitude of the selected tap of the channel response. Some improvements are proposed in [18] to find better constraints. Lee [13] mentions the possible use of the CM or Godard cost function as a constraint. A general equalization method for Multiple Input Multiple Output (MIMO) channels using the Godard or CM cost function is proposed in [19]. The CMA MUD methods in this dissertation will be based on the recent work done in [19–23].

The third kind of blind MUD detection methods is based on linear prediction methods [24], or linear prediction like methods [25]. The main idea of the linear prediction approach is to use the null subspace of the desired user's spreading code matrix to estimate the channel and then to estimate the detector. One possible drawback of the linear prediction method, is that the channel estimation may suffer from system noise and computation errors [13], which will deteriorate the symbol detection. In [26], it is shown that a direct blind equalizer can be obtained by using linear prediction to estimate the column vector subspace of the channel without estimating the channel itself. Instead of two stages of linear prediction in [26], only one stage is required for CDMA [13].

1.4 APPLYING THE CONSTANT MODULUS CRITERION TO MUD

The most studied and implemented adaptation algorithm of the 1990s is the CMA [4]. The CMA is a special case of the Goddard algorithm [27,28]. The CMA seeks to minimize a cost defined by the CM criterion. This criterion penalizes deviations in the modulus (magnitude) of the equalized signal away from a fixed value. A major advantage of the CMA is that it is a blind algorithm, and does not need

a known training sequence to be transmitted. It is obvious that the inclusion of such a training signal sacrifices valuable channel resources. The CMA is a stochastic gradient algorithm [27], which applies a memoryless non-linearity at the output of the linear FIR equalizer in order to generate the desired response with each iteration. The nature of the non-linearity in the CMA will be discussed later in the dissertation. The CM and MSE criteria have several similarities with regard to their cost surfaces. The cost surface is a multi-dimensional surface of the MSE or CM cost versus the equalizer coefficients. Under AWGN conditions, MSE and CM cost functions also have exactly the same minima [4].

The CMA was widely implemented in a Single Input Single Output (SISO) channel for adaptive equalization. The CDMA channel is a MIMO channel, as several users share the same bandwidth. Tugnait [19] proposed the use of the Goddard cost function in a MIMO channel. In this way the CMA can be generalized to the MIMO case, and can specifically be applied to MUD in CDMA systems. The use of a linear constraint, utilizing information about the desired user signature vector, can prevent the detector from locking on to interfering user signals. After this, several authors proposed the use, and evaluated the use of the linearly constrained constant modulus detector. These issues are discussed in detail in Chapter 5. A notable advantage of the linearly constrained constant modulus (LCCM) detector is the fact that it requires no more information than the SUD. It only requires knowledge of the timing and the signature waveform of the desired user. In this dissertation, variants of the LCCM detector i.e. the LCCM and the linearly constrained differential constant modulus (LCDCM) detectors will be thoroughly studied and evaluated.

1.5 CONTRIBUTIONS OF THIS DISSERTATION

This dissertation focuses on the application of the constant modulus algorithm to the multiuser detection problem within the context of adaptive linear detector structures. Two different forms of the constant modulus multiuser detector are analyzed, evaluated and compared with each other and the MMSE detector. The main unique contributions of this dissertation can be itemized as follows:

- Existing signal, channel and detector models are expanded to encompass the complex valued *multipath* DS-CDMA channels.
- For the first time, a global convexity condition is extensively derived for the LCCM detector cost function.
- Simulation results for different channel types are generated and discussed. These channel types range from the additive white Gaussian noise (AWGN) channel to multipath fading channels.
- The application of the variants of the LCCM detector to non-linear multipath fading channels are thoroughly investigated. The issues and limitations with respect to non-linear channels are

discussed, analyzed and evaluated.

The following *research outputs* were generated during the completion of this dissertation:

1. Submission of a paper to IEEE Transactions on Communications, June 2002 [29].
2. International Conference presentation at IEEE Africon October 2002, George, South Africa [30].
3. Submission of an abstract for a paper in a special issue of the Transactions of the SAIEE.

The following conclusions are presented as determined by analysis and simulation:

- The blind adaptive LCCM detector suffers from ill-convergence under the condition that the desired user amplitude falls below a certain level. This is due to the LCCM cost function exhibiting undesired minima under this condition.
- The blind adaptive LCDCM detector converges independent of desired user amplitude. The LCDCM cost function exhibits a global minimum in an AWGN channel.
- Tap weight vector convergence of both LCCM and LCDCM detectors approach the mean tap weight vector of the MMSE detector in an AWGN channel, but not in a multipath fading channel.
- Under normal operation, the MMSE detector can combine static multipaths, while the LCCM and LCDCM detectors attempt to cancel it out.
- If inverse channel plus noise estimation can be used within the linear constraint, the LCCM and LCDCM detectors can effectively combine the multiple paths, providing that the multipath channel plus noise inverse can be accurately modelled within the length of the detector.
- BER and SIR simulation measurements show that the blind LCCM and LCDCM detectors exhibit similar performance to that of the non-blind MMSE detector in an AWGN channel. Ill convergence of the LCCM detector is demonstrated if the desired user amplitude falls below a certain level as calculated analytically.
- BER simulation measurements show effective operation of the LCCM and LCDCM detectors in a minimum phase *non-fading* multipath channel.
- In a multipath fading channel, the resulting channel may at times be ill-behaved. This means that the inverse channel plus noise may not accurately modelled by a linear filter. Consequently, the linear MMSE, LCCM and LCDCM detectors are at times unable to equalize the channel, and the eye closes.

1.6 OUTLINE OF DISSERTATION CHAPTERS

This chapter gives a qualitative introduction to the CDMA environment and a heuristic view of the multiuser detection problem. Application of the CMA to multiuser detection is briefly visited, and a general problem definition is stated. Chapter 2 introduces the reader to the CDMA signal and channel models that will be extensively used within the dissertation. Chapter 3 makes the reader intuitively aware of the issues regarding the multiuser detection problem. Two user graphical examples are used to assist in this process. Several criteria are given whereby multiuser detectors may be evaluated. Chapter 4 starts with a generalized discussion on all linear multiuser detection techniques. The MMSE detector model is presented in detail, supported by a rigorous theoretical performance evaluation. Chapter 5 treats the linearly constrained constant modulus detector and an improved variant, the linearly constrained differential constant modulus detector. The cost criteria of these detectors are analytically analyzed and scrutinized. The advantages and disadvantages of both detectors are discussed. In Chapter 6, simulation results are presented. These results, along with the theoretical results obtained in Chapters 4 and 5 are comparatively discussed. Conclusions are drawn, and areas for further possible study and investigation are also proposed. The dissertation outline is depicted in Figure 1.2.

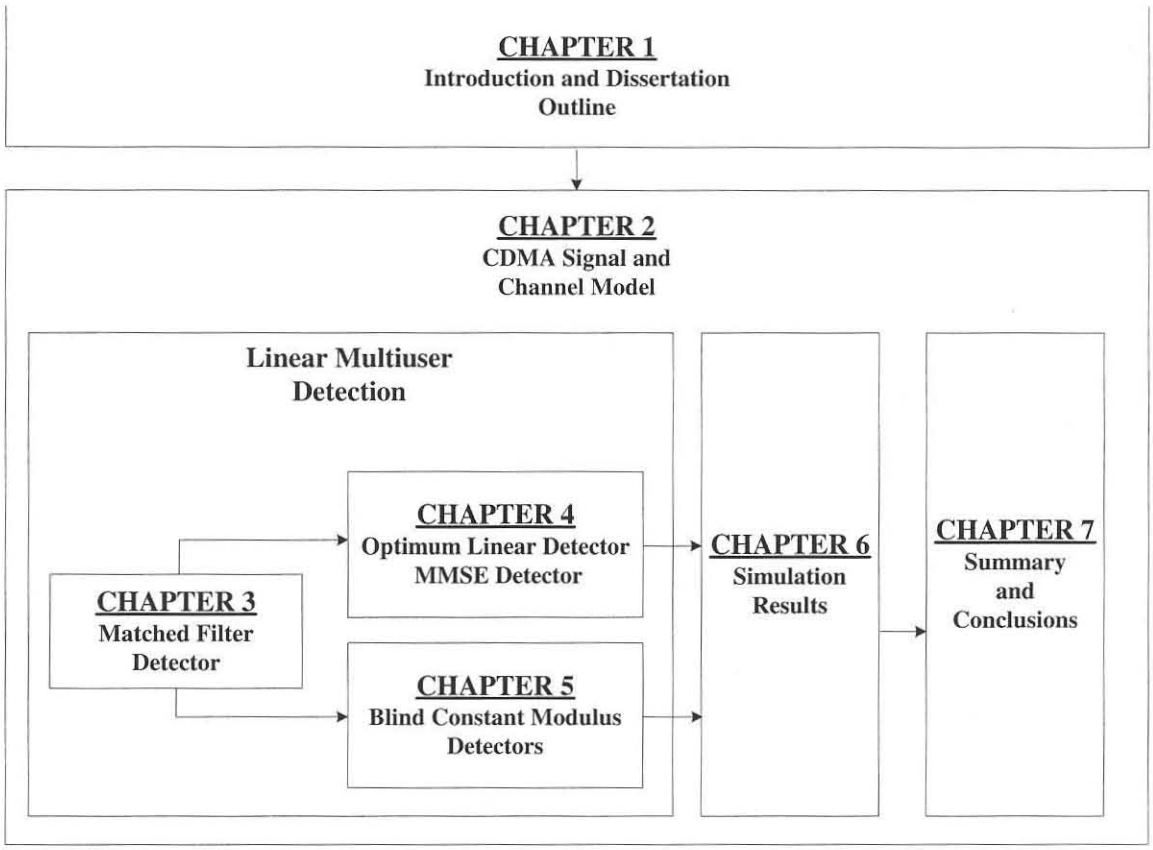

Figure 1.2: Graphical representation of the structure and outline of the dissertation.

CHAPTER TWO

SYNCHRONOUS AND ASYNCHRONOUS CDMA MODELS

This chapter contains the mathematical signal and channel models that are to be used to analyze multiuser detection methods. We will limit ourselves to the baseband case for simplicity. The analysis contained here is largely based on the approach followed by Verdu in [31] and Rappaport in [32].

2.1 THE CDMA SIGNAL MODEL

Consider a CDMA channel that is shared by K simultaneous users. Each user is assigned a signature waveform. For user k, the waveform is denoted by

$$s_k(t) = \sum_{n=0}^{N-1} a_k(n)\, p(t - nT_c), \quad 0 \le t \le T \tag{2.1}$$

where $\{a_k(n), 0 \le n \le N-1\}$ is a pseudo-noise sequence, consisting of N chips that take the values $\{\pm 1\}$ and $p(t)$ is a pulse of duration T_c, where T_c is a chip interval. Without loss of generality, we assume that all K signature waveforms have unit energy, i.e.

$$\|s_k(t)\| = \int_0^T s_k(t)dt = 1. \tag{2.2}$$

The cross correlations (or inner products) between pairs of signature waveforms play an important role in the metrics for the signal detector and on its performance. We define the cross correlations between two arbitrary signature waveforms for the synchronous case.

$$\rho_{kj} = \langle s_k, s_j \rangle = \int_0^T s_k(t)s_j(t)dt \tag{2.3}$$

Note that by the Cauchy-Schwartz inequality and (2.2) we have

$$|\rho_{kj}| = \langle s_k, s_j \rangle \leq \|s_k\| \, \|s_j\| = 1. \tag{2.4}$$

Let us also define the cross correlation matrix,

$$\mathbf{R} = \{\rho_{kj}\} \tag{2.5}$$

which has diagonal elements equal to one and is symmetric nonnegative definite, because for any K-vector $\mathbf{c} = (c_1, \ldots, c_K)^T$ we have

$$\mathbf{c}^T \mathbf{R} \mathbf{c} = \left\| \sum_{k=1}^{K} c_k s_k \right\|^2 \geq 0. \tag{2.6}$$

Therefore the cross correlation matrix $\mathbf{R}$ is positive definite if and only if the signature waveforms $\{s_1, \ldots, s_K\}$ are linearly independent.

Concerning the asynchronous case, Figure 2.1 shows a schematic representation of the cross correlation between two synchronous users.

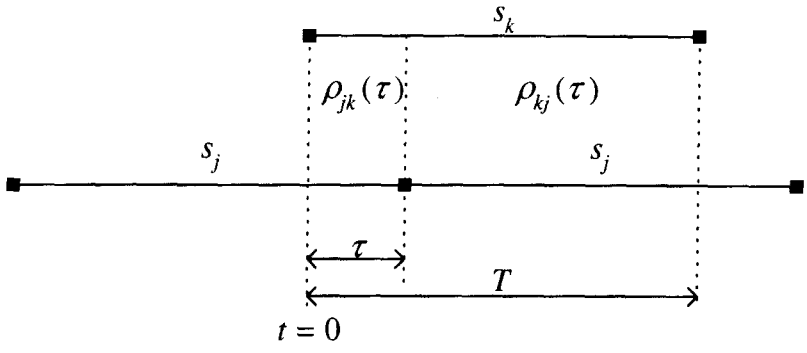

Figure 2.1: Schematic representation of the cross correlation between two synchronous users

As can be seen in Figure 2.1, we must define two cross correlations between every pair of signature waveforms that depends on τ, the offset between the two signatures. If $k < j$, we write the cross correlations as

$$\rho_{kj}(\tau) = \int_{\tau}^{T} s_k(t) s_j(t - \tau) dt \tag{2.7}$$

$$\rho_{jk}(\tau) = \int_{0}^{\tau} s_k(t) s_j(t + T - \tau) dt \tag{2.8}$$

where $t \in [0, T]$, and T denoted the signature waveform length in seconds.

2.2 DISCRETE-TIME SYNCHRONOUS MODELS

Multiuser detectors commonly have a front end which has the task of obtaining a discrete time process from a received continuous waveform $y(t)$. Generally, continuous to discrete conversion can be done by correlating $y(t)$ with deterministic signals. In communication theory, there are two types of deterministic signals of interest. These are matched signature waveforms (matched filters) and orthonormal signals.

The basic K-user CDMA model, consisting of the sum of antipodally modulated synchronous signature waveforms embedded in AWGN is given by

$$y(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + \sigma n(t), \qquad t \in [0, T] \tag{2.9}$$

where A_k is the received signal amplitude of the kth user, $b_k \in \{\pm 1\}$ is the bit transmitted by the kth user, $s_k(t)$ is the deterministic signature waveform of user k, $n(t)$ is the white Gaussian noise component with unit power spectral density, and σ the noise variance.

2.2.1 MATCHED FILTER OUTPUTS

Using equations (2.3) and (2.9), we can express the matched filter output of the kth user as

$$y_k = A_k b_k + \sum_{j \neq k} A_j b_j \rho_{jk} + n_k \tag{2.10}$$

where

$$n_k = \sigma \int_0^T n(t) s_k(t) dt \tag{2.11}$$

is a Gaussian random variable with zero mean and variance equal to σ^2, since by (2.2), $s_k(t)$ has unit energy. We refer to n_k as the noise component of user k.

If we express (2.10) in vector matrix notation, we obtain

$$\mathbf{y} = \mathbf{RAb} + \mathbf{n} \tag{2.12}$$

where $\mathbf{R}$ is the normalized cross correlation matrix, $\mathbf{y} = [y_1, \ldots, y_K]^T$, $\mathbf{b} = [b_1, \ldots, b_K]^T$ and $\mathbf{A} = \text{diag}\{A_1, \ldots, A_K\}^T$. The vector $\mathbf{n}$ is a zero mean Gaussian random vector with a covariance matrix equal to

$$E[\mathbf{n}\mathbf{n}^T] = \sigma^2 \mathbf{R} \tag{2.13}$$

It will later be shown that no information relevant to demodulation is lost by the bank of matched filters. This means that $y(t)$ can be replaced with $\mathbf{y}$ without loss of optimality.

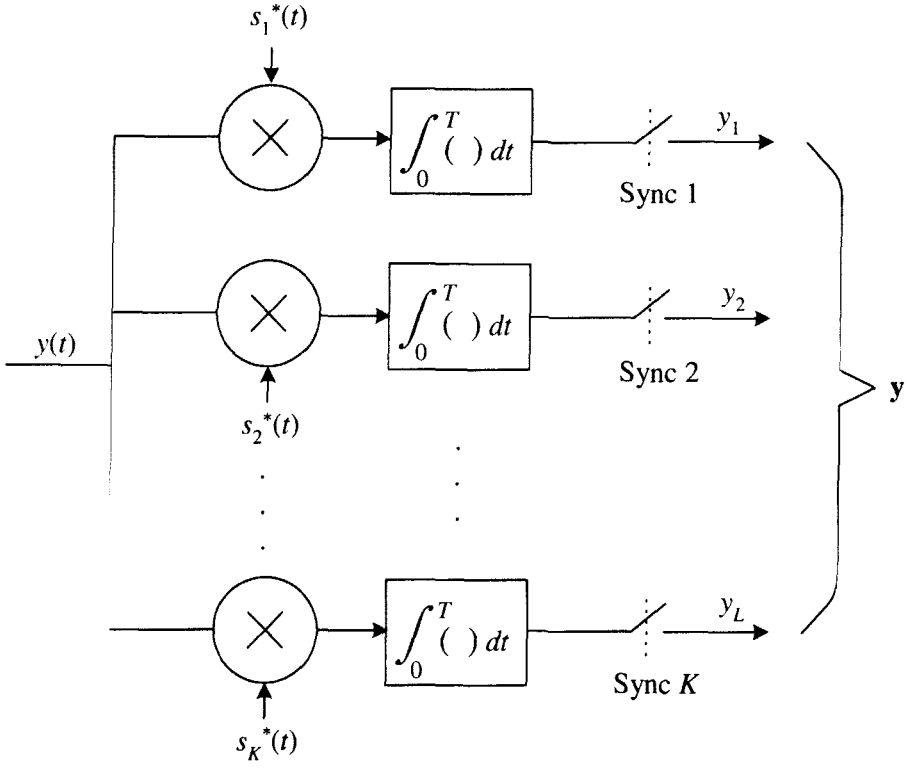

Figure 2.2: Block diagram illustration of the complex matched filter receiver

The unnormalized cross correlation matrix whose (j,k) elements is given by $\langle A_j s_j, A_k s_k \rangle$, is written as

$$\mathbf{H} = \mathbf{ARA} \tag{2.14}$$

When the receiver front end consists of a bank of matched filters, we have seen that we can replace the model in (2.9) with the linear Gaussian vector matrix model in (2.12). The same model can be generalized to encompass complex numbers. The only difference is that the output of the matched filter is given by

$$y_k = \langle y, s_k^* \rangle = \int_0^T y(t) s_k^*(t) dt \tag{2.15}$$

where * denotes the complex conjugate. This means that the cross correlation values are given by

$$\rho_{kj} = \int_0^T s_k^*(t) s_j(t) dt \tag{2.16}$$

yielding the same model as in (2.12) encompassing complex values,

$$\mathbf{y} = \mathbf{RAb} + \mathbf{n} \tag{2.17}$$

where the correlation matrix $\mathbf{R}$ is in Hermitian form. $\mathbf{A}$ is a complex diagonal matrix and n is a complex valued Gaussian vector with independent real and imaginary components and with a covariance matrix equal to $2\sigma^2\mathbf{R}$.

In the complex valued asynchronous case, the cross correlation values are given by

$$\rho_{kj}(\tau) = \int_{\tau}^{T} s_k(t)s_j^*(t-\tau)dt \tag{2.18}$$

and

$$\rho_{jk}(\tau) = \int_{0}^{\tau} s_k(t)s_j^*(t+T-\tau)dt. \tag{2.19}$$

2.2.2 WHITENED MATCHED FILTER MODEL

Notice that the noise between users is correlated in the standard discrete time synchronous model. This causes difficulty in the evaluation of performance of the various multiuser detection techniques. We can correct this with the use of a whitening filter as described below.

Proposition 2.1 *(Cholesky Factorization) For every positive definite Hermitian matrix* $\mathbf{R}$*, there exists a unique lower triangular matrix* $\mathbf{F}$ *(i.e.* $F_{ik} = 0$ *for* $i < k$*) with positive diagonal elements such that*

$$\mathbf{R} = \mathbf{F}^H\mathbf{F}$$

where $\mathbf{F}^H$ *denotes the Hermitian (complex conjugate) transpose of F.*

For brevity we shall denote the inverse of a Hermitian transpose of a matrix by

$$\left(\mathbf{F}^H\right)^{-1} \stackrel{\text{def}}{=} \mathbf{F}^{-H} \tag{2.20}$$

If the matched filter outputs $\mathbf{y}$ are processed by the matrix $\mathbf{F}^{-H}$, called a whitening filter, we obtain the whitened matched filter model

$$\begin{aligned} \bar{\mathbf{y}} &= \mathbf{F}^{-H}\mathbf{y} \\ &= \mathbf{F}^{-H}\mathbf{F}^H\mathbf{F}\mathbf{A}\mathbf{b} + \mathbf{F}^{-H}\mathbf{n} \\ &= \mathbf{F}\mathbf{A}\mathbf{b} + \bar{\mathbf{n}} \end{aligned} \tag{2.21}$$

where $\bar{y}_k$ contains contributions for users $1 \ldots k$, but not from users $k+1 \ldots K$. The covariance matrix of $\bar{\mathbf{n}}$ is

$$\begin{aligned} E[\bar{\mathbf{n}}\bar{\mathbf{n}}^H] &= 2\sigma^2\mathbf{F}^{-H}\mathbf{R}\mathbf{F}^{-1} \\ &= 2\sigma^2\mathbf{F}^{-H}\mathbf{F}^H\mathbf{F}\mathbf{F}^{-1} \\ &= 2\sigma^2\mathbf{I} \end{aligned} \tag{2.22}$$

where $\mathbf{I}$ is the identity matrix. As the name suggests, the whitened matched filter causes the noise components to be independent as in (2.22).

2.2.3 ORTHONORMAL PROJECTIONS

In the previous two models, the dimensionality of the vectors in (2.12) and (2.17) is equal to the number of users. In some situations (such as when the signature waveforms of some interferers are unknown) other models (with possibly different dimensionality) are useful. A receiver utilizing orthonormal projections is termed a correlation receiver (Figure 2.3).

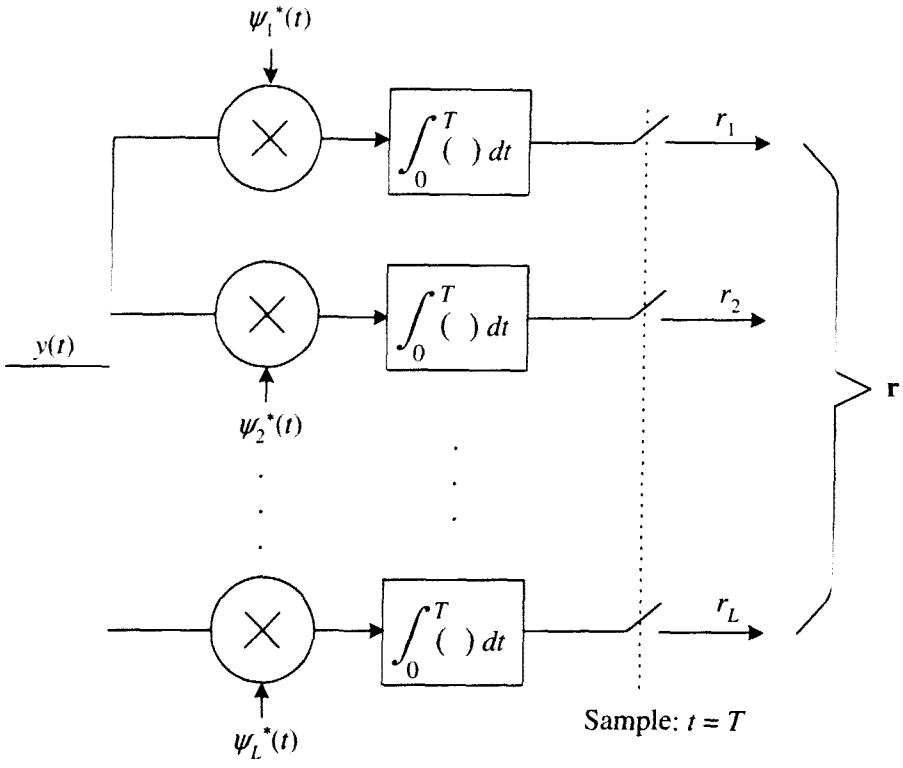

Figure 2.3: Block diagram description of the orthonormal projection correlation receiver

Let $\{\psi_1, \ldots, \psi_L\}$ be a set of L complex orthonormal signals defined on $[0, T]$. The complex *signature vector* $\mathbf{s_k}$ of the kth user is the L dimensional representation of s_k on the basis $\{\psi_1, \ldots, \psi_L\}$. That is to say the lth component of the column vector $\mathbf{s}_k$ is

$$s_{kl} = \int_0^T s_k(t)\psi_l^*(t)dt \tag{2.23}$$

Furthermore, we define the lth component of the vector $\mathbf{r}$ as

$$r_l = \int_0^T y(t)\psi_l^*(t)dt \tag{2.24}$$

The column vector can then be written as

$$\begin{aligned} \mathbf{r} &= \sum_{k=1}^{K} A_k b_k \mathbf{s}_k + \sigma \mathbf{m} \\ &= \mathbf{S}\mathbf{A}\mathbf{b} + \sigma \mathbf{m} \end{aligned} \tag{2.25}$$

where $\mathbf{m}$ is an L dimensional complex Gaussian vector with independent unit variance components. Now we introduce a $L \times K$ matrix of complex signature vectors

$$\begin{aligned} \mathbf{S} &= \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_K \end{bmatrix} \\ &= \begin{bmatrix} s_{11} & \cdots & s_{K1} \\ \vdots & \ddots & \vdots \\ s_{1L} & \cdots & s_{KL} \end{bmatrix} \end{aligned} \tag{2.26}$$

The bits of the different users are uncorrelated, resulting in a covariance matrix equal to

$$\begin{aligned} E[\mathbf{r}\mathbf{r}^H] &= 2\sigma^2 \mathbf{I} + 2 \sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H \\ &= 2\sigma^2 \mathbf{I} + 2\mathbf{S}\mathbf{A}^2\mathbf{S}^H \end{aligned} \tag{2.27}$$

The finite dimensional model in (2.25) holds regardless of whether the L orthonormal signals $\{\psi_1, \ldots, \psi_L\}$ span the signature waveforms $\{s_1, \ldots, s_K\}$. An example of a set of orthonormal signals that span the signature waveforms is a DS-CDMA system where L is equal to the number of chips per symbol and the orthonormal signals are the delayed chip waveforms $\psi_i = p(t - (i-1)T_c)$.

If the signature waveforms are spanned by $\{\psi_1, \ldots, \psi_L\}$, then the $K \times K$ cross correlation matrix simply becomes

$$\mathbf{R} = \mathbf{S}^H \mathbf{S} \tag{2.28}$$

Furthermore

$$\|\mathbf{s}_k\| = 1 \tag{2.29}$$

and all the information contained in $\mathbf{y}$ is also contained in $\mathbf{r}$, because the matched filter outputs can be expressed as a linear combination of the components of $\mathbf{r}$, i.e.

$$\mathbf{y} = \mathbf{S}^H \mathbf{r} \tag{2.30}$$

2.3 DISCRETE-TIME ASYNCHRONOUS MODELS

For a simplified notation, we shall label the users chronologically. We assume without loss of generality that $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_K$. If we generalize (2.9) to the complex asynchronous case, the complex asynchronous CDMA model becomes

$$y(t) = \sum_{k=1}^{K} \sum_{i=1}^{M} A_k b_k[i] s(t - iT - \tau_k) + \sigma n(t) \tag{2.31}$$

taking into account that the users send a complex bitstream $b_k[-M], \ldots, b_k[0], \ldots, b_k[M]$. The length of the packets transmitted by each user is assumed to be equal to $(2M + 1)$.

In the context of this model, users initiate and terminate their transmissions within T time units from each other. This presupposes some form of block synchronism, if not symbol synchronism. This assumption allows us to focus on the offsets modulo T, and does not impact on the generality of the analysis, because of typically large values of M. Figure 2.4 shows the symbol epochs for three asynchronous users in a case where $M = 1$.

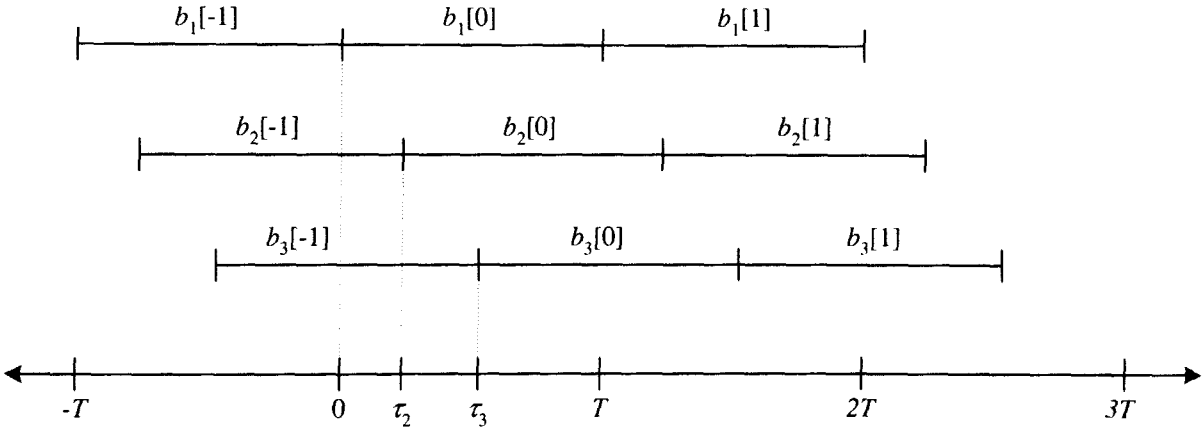

Figure 2.4: Schematic representation of the symbol epochs for three users if $M = 1$

The synchronous channel corresponds to the special case of (2.31) where $\tau_1 = \tau_2 = \ldots = \tau_K = 0$.

2.3.1 INTERSYMBOL INTERFERENCE

Consider the special case in which all the complex received amplitudes and all the complex signature waveforms are equal, i.e. $A_1 = A_2 = \ldots = A_K$ and $s_1 = s_2 = \ldots = s_K$, and in which the offsets satisfy

$$\tau_k = \frac{(k-1)T}{K}. \tag{2.32}$$

The asynchronous model then becomes

$$\begin{aligned} y(t) &= \sum_{k=1}^{K}\sum_{i=-M}^{M} Ab_k[i]s\left(t - iT - \frac{(k-1)T}{K}\right) + \sigma n(t) \\ &= \sum_{j} Ab[j]s\left(t - \frac{jT}{K}\right) + \sigma n(t) \end{aligned} \tag{2.33}$$

where we have denoted $b[iK + k - 1] = b_k[i]$. The channel in (2.33) is, in fact, the single user white Gaussian channel with intersymbol interference. ISI is a phenomenon encountered in both synchronous and asynchronous CDMA systems. It may be due to a frequency selective (discrete multipath) channel or partial response signalling (to increase the signature time bandwidth product). We will discuss the frequency selective (or multipath) channel later in this chapter, as it is commonly encountered in the mobile channel. We will also derive the discrete multipath channel from the continuous time dispersion channel filter model.

2.3.2 ASYNCHRONOUS VECTOR MATRIX MODEL

When using (2.31) with (2.18) and (2.19), the matched filter outputs can be expressed as

$$\begin{aligned} y_k[i] &= A_k b_k[i] \\ &\quad + \sum_{j<k} A_j b_j[i+1]\rho_{kj} + \sum_{j<k} A_j b_j[i]\rho_{jk} \\ &\quad + \sum_{j>k} A_j b_j[i]\rho_{kj} + \sum_{j>k} A_j b_j[i-1]\rho_{jk} \\ &\quad + n_k[i] \end{aligned} \tag{2.34}$$

where

$$n_k[i] = \sigma \int_{\tau_k+iT}^{\tau_k+iT+T} n(t)s_k^*(t - iT - \tau_k)dt. \tag{2.35}$$

The first line of equation (2.34) is the desired information. The second line is the interference due to earlier users and the third line represents the interference due to later users. We can write equation (2.34) in matrix form,

$$\begin{aligned} \mathbf{y}[i] &= \mathbf{R}^H[1]\mathbf{A}\mathbf{b}[i+1] + \mathbf{R}[0]\mathbf{A}\mathbf{b}[i] \\ &\quad + \mathbf{R}[1]\mathbf{A}\mathbf{b}[i-1] + \mathbf{n}[i] \end{aligned} \tag{2.36}$$

where the zero mean Gaussian process $\mathbf{n}[i]$ has the autocorrelation matrix

$$E[\mathbf{n}[i]\mathbf{n}^H[j]] = \begin{cases} 2\sigma^2\mathbf{R}^H[1] & \text{if} j = i+1 \\ 2\sigma^2\mathbf{R}[0] & \text{if} j = i \\ 2\sigma^2\mathbf{R}[1] & \text{if} j = i-1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.37}$$

and the complex valued matrices $\mathbf{R}[0]$ and $\mathbf{R}[1]$ are defined by

$$R_{jk}[0] = \begin{cases} 1 & \text{if} j = k \\ \rho_{jk} & \text{if} j < k \\ \rho_{kj} & \text{if} j > k \end{cases} \tag{2.38}$$

and

$$R_{jk}[1] = \begin{cases} 0 & \text{if} j \geq k \\ \rho_{kj} & \text{if} j < k \end{cases}. \tag{2.39}$$

For example, in the three user case

$$\mathbf{R}[0] = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}, \tag{2.40}$$

$$\mathbf{R}[1] = \begin{bmatrix} 0 & \rho_{21} & \rho_{31} \\ 0 & 0 & \rho_{32} \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.41}$$

The vector matrix discrete time model in (2.36) can be represented in the z-transform domain

$$\mathbf{S}(z) = \mathbf{R}^H[1]z + \mathbf{R}[0] + \mathbf{R}[1]z^{-1}. \tag{2.42}$$

This means we can also represent (2.42) as the combined asynchronous correlation matrix

$$\mathbf{R}_a = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}^H[1] & 0 & \cdots & 0 \\ \mathbf{R}[1] & \mathbf{R}[0] & \mathbf{R}^H[1] & \cdots & \vdots \\ 0 & \mathbf{R}[1] & \ddots & \vdots & 0 \\ \vdots & \vdots & \cdots & \mathbf{R}[0] & \mathbf{R}^H[1] \\ 0 & \cdots & 0 & \mathbf{R}[1] & \mathbf{R}[0] \end{bmatrix}. \tag{2.43}$$

The z domain model is depicted in Figure 2.5, where $\bar{\mathbf{n}}[i]$ is independent Gaussian with covariance matrix $2\sigma^2\mathbf{I}$.

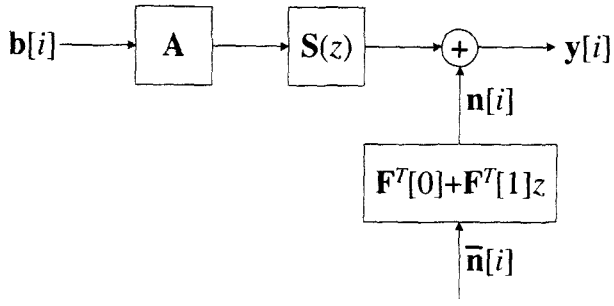

Figure 2.5: Block diagram of the z domain vector matrix model of equations (2.36) and (2.42)

Note that if the signature waveforms have a duration larger than T, then the model has to be generalized to incorporate crosscorrelation matrices $\mathbf{R}[2], \ldots, \mathbf{R}[L]$, where L is the length of the intersymbol interference. The choice of the $K \times K$ matrices $\mathbf{F}[0]$ and $\mathbf{F}[1]$ in Figure 2.5 is governed by the following proposition:

Proposition 2.2 *(z Transform Model Cholesky Factorization) The complex valued matrix* $\mathbf{S}[z]$ *in (2.42) can be expressed as*

$$\mathbf{S}(z) = [\mathbf{F}[0] + \mathbf{F}[1]z]^H \, [\mathbf{F}[0] + \mathbf{F}[1]z^{-1}] \tag{2.44}$$

where $\mathbf{F}[0]$ *is lower triangular and* $\mathbf{F}[1]$ *is upper triangular with zero diagonal such that*

$$\mathbf{R}[0] = \mathbf{F}^H[0]\mathbf{F}[0] + \mathbf{F}^H[1]\mathbf{F}[1] \tag{2.45}$$

$$\mathbf{R}[1] = \mathbf{F}^H[0]\mathbf{F}[1] \tag{2.46}$$

$$\det \mathbf{F}[0] = \exp\left(\frac{1}{2}\int_0^1 \log\left(\det \mathbf{S}\left(e^{j2\pi f}\right)\right) df\right) \tag{2.47}$$

Furthermore, if $\det \mathbf{S}(e^{j2\omega}) > 0$ *for all* $\omega \in [-\pi, \pi]$, *then* $[\mathbf{F}[0] + \mathbf{F}[1]z^{-1}]^{-1}$ *is causal and stable.*

As with (2.21), if the vector sequence of matched filter outputs is fed into the filter $[\mathbf{F}[0]+\mathbf{F}[1]z^{-1}]^{-1}$, the output sequence is given by

$$\bar{\mathbf{y}}[i] = \mathbf{F}[0]\mathbf{A}\mathbf{b}[i] + \mathbf{F}[1]\mathbf{A}\mathbf{b}[i-1] + \bar{\mathbf{n}}[i] \tag{2.48}$$

where as with (2.22), $\bar{\mathbf{n}}[i]$ is independent Gaussian with covariance matrix $2\sigma^2\mathbf{I}$.

As with the synchronous case, alternative finite dimensional models can be used with a set of orthonormal waveforms that span the signature space, i.e. all the signature waveforms and their delayed

versions. In a direct sequence spread spectrum system, this can be accomplished by chip matched filters sampled at the chip rate times the number of users. Nevertheless, for approximately band-limited chip waveforms, it is sufficient to sample at the Nyquist rate.

2.4 THE FADING MOBILE CHANNEL MODEL

We will now consider mobile channel models for the evaluation of multiuser detection methods and CDMA transmission. Mobile channels are dominated by a phenomenon called *fading*. Fading is the variation in signal strength over a period of time. We will mostly concern ourselves with *small scale* fading. Small scale fading is rapid signal strength variation over time or distance. We will assume the large scale fading (due to shadowing) to be quasi-stationary, and thus less relevant to our comparative evaluation of multiuser detection schemes.

There are two main types of small scale fading in a mobile channel. The first is fading due to multipath, and the other is fading due to Doppler spread. Multipath delay causes time dispersion and frequency selective fading, while Doppler spread causes frequency dispersion and time selective fading. We can subdivide multipath fading into two more components.

Flat Fading - In this case, the bandwidth of the signal is smaller than the bandwidth of the channel. This also means that the delay spread is smaller than the symbol period. The spectral characteristics of the transmitted signal is preserved at the receiver, thus no inter symbol interference (ISI) is introduced.

Frequency Selective Fading - Here, the bandwidth of the signal is greater than the bandwidth of the channel. Furthermore, the delay spread is greater than the symbol period. Frequency selective fading introduces time dispersion between the symbols, introducing ISI.

Fading based on Doppler spread, can also be subdivided into two categories. These are fast fading and slow fading.

Fast Fading - In this case, the channel has a large Doppler spread. Furthermore, the coherence time is smaller than the symbol period. Here, the channel variations occur faster that the baseband signal variations.

Slow Fading - In contrast with fast fading, the channel has a small Doppler spread. This means that the coherence time is greater than the symbol period, and channel variations appear to be slower than baseband signal variations.

Another mobile channel effect is Doppler shift due to the relative motion between transmitter and receiver. This is a carrier frequency and velocity dependent frequency offset on each of the multipath components. This effect is taken into account in Clarke's model, which we will discuss now.

2.4.1 RAYLEIGH FADING DUE TO DOPPLER SPREAD - CLARKE'S MODEL

Clarke developed a model to deduce the statistical characteristics from the scattered electromagnetic fields of the received signal at the mobile receiver [33]. In this model, the envelope of the received E-field E_z is the square root of the sum of two squared Gaussian random variables. By random variable transformation, we have that the received signal envelope of a certain propagation path has a Rayleigh distribution given by

$$f_R(r) = \begin{cases} r^2 \exp\left(\frac{-r^2}{2}\right), & 0 \leq r \leq \infty. \\ 0, & r < 0. \end{cases} \tag{2.49}$$

To evaluate the probability of error of a CDMA detector (or any other digital communication detector) in a Rayleigh fading channel, the signal-to-noise ratio γ must be averaged over all possible fading signal amplitudes. That is to say

$$P_{e,R} = \int_0^\infty P_e(\gamma) p(\gamma) d\gamma, \tag{2.50}$$

where $P_e(\gamma)$ is the Gaussian channel error probability for an arbitrary modulation at a specific value of signal-to-noise ratio γ, and $P_{e,R}$ is the error probability for the Rayleigh faded signal.

2.4.2 MULTIPATH TIME DISPERSION MODEL

The mobile channel can be modelled as a linear filter. This means that the small scale variations of the mobile radio channel can be characterized by the impulse response of the mobile channel. The impulse response model of the mobile channel is useful, since it may be used to predict and compare the performances of many mobile communication systems under many different conditions.

To show that the mobile channel can be modelled as a linear filter with a time varying response, consider the case in Figure 2.6 where the time variation is only due to the motion of the mobile. We assume that receiver moves along the ground at some constant speed v. For a fixed position d, the channel between the receiver and the transmitter can be modelled as a linear time invariant system.

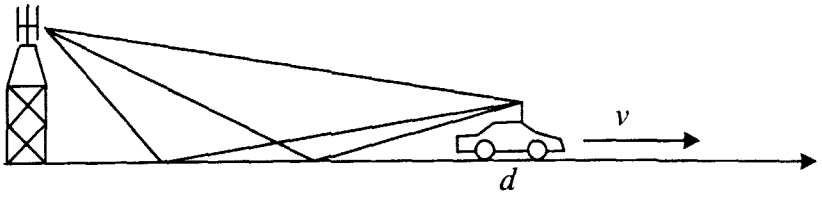

Figure 2.6: The mobile radio channel as a function of time and space.

Due to different multipath waves that differ from position to position, the impulse response of the linear time invariant system should be a function of the position of the receiver. Let the channel impulse response of user k be depicted by $h_k(d,t)$. The effect of time dispersion on the basic CDMA model is that the signature waveform seen at the receiver is the complex convolution

$$\tilde{s}_k(d,t) = s_k(t) \otimes h_k(d,t) = \int_{-\infty}^{\infty} s_k(\tau) h_k(d, t-\tau) d\tau. \tag{2.51}$$

For a causal system, $h_k(d,t) = 0$ for $t < 0$. This results in the following equation

$$\tilde{s}_k(d,t) = \int_{0}^{t} s_k(\tau) h_k(d, t-\tau) d\tau. \tag{2.52}$$

If the receiver moves at a constant speed v, the position of the receiver can be expressed as

$$d = vt. \tag{2.53}$$

Substituting (2.53) in equation (2.52), we obtain

$$\tilde{s}_k(vt,t) = \int_{0}^{t} s_k(\tau) h_k(vt, t-\tau) d\tau \tag{2.54}$$

We assume v to be constant with respect to symbol time T, that is $\tilde{s}(vt,t)$ is only a function of t. Therefore, (2.53) can be written as

$$\tilde{s}(vt,t) = \int_{0}^{t} s_k(\tau) h_k(vt, t-\tau) d\tau = s_k(t) \otimes h_k(vt,t) = s_k(t) \otimes h_k(d,t) \tag{2.55}$$

The kth user impulse response $h_k(t,\tau)$ completely characterizes the channel as a function of both t and τ. The variable t represents the time variations due to motion, and τ represents the channel multipath delay for a fixed value of t. The output of the "channel filter" for user k is given by

$$\tilde{s}_k(t) = \int_{0}^{t} s_k(\tau) h_k(t,\tau) d\tau = s_k(t) \otimes h_k(t,\tau) \tag{2.56}$$

2.4.2.1 DISCRETE TIME CHANNEL IMPULSE RESPONSE

We can divide the multipath delay axis τ of every user's impulse response into discrete bins. These equal time delay segments are called *excess delay* bins. Each bin has a time delay width of $\tau_{p+1} - \tau_p$, where τ_0 is equal to the time instant of the first arriving signal and equal to τ_k. The first bin from τ_0 to τ_1 has a bin width of $\Delta\tau$, as with all the other bins. This means that $\tau_0 = 0$, $\tau_1 = \Delta\tau$, and $\tau_p = p\Delta\tau$, for $p = 0$ to $p = P-1$, where P represents the total number of equally spaced multipath components. The size of $\Delta\tau$ determines the time delay resolution of the channel model. The useful frequency span of the model is shown to be $1/(2\Delta\tau)$. This means that signals of maximum bandwidth $1/(2\Delta\tau)$ can be evaluated using this model. We assume that the delay resolution is equal for all users. There are terms that apply to this model that have to be briefly discussed. The first is *excess delay*, being the

relative delay of the pth multipath component to the first arriving component, and is denoted by τ_p. The *maximum excess delay* is given by $P\Delta\tau$.

The received signal consists of a series of attenuated, time delayed and phase shifted versions of the transmitted signals. The impulse response of user k can be expressed as

$$h_k(t,\tau) = \sum_{p=0}^{P-1} A_{k,p}(t,\tau) \exp\left[j(2\pi f_{d_k}\tau_p + \phi_{k,p}(t,\tau))\right] \delta(\tau - \tau_p) \tag{2.57}$$

where $A_{k,p}(t,\tau)$ and τ_p are the real amplitudes at time t and excess delays, respectively, of the pth multipath component of user k. The phase term $(2\pi f_{d_k}\tau_p + \phi_{k,p}(t,\tau))$ represents the phase shift, due to a doppler shift f_{d_k} and other channel effects (such as Rayleigh fading), of the pth multipath component at time t of user k. The phase term can be combined and represented by the term $\theta_{k,p}(t,\tau)$. Figure 2.7 is a graphical representation of equation (2.57) for a certain channel at different times t. Note that some of the excess delay bins (or multipath components) may have a zero amplitude.

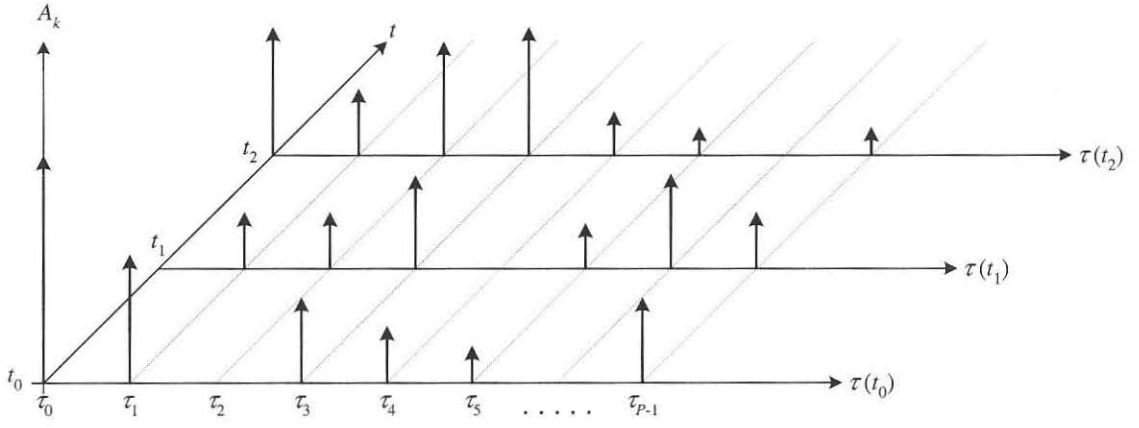

Figure 2.7: The time varying discrete time impulse response model for a specific multipath radio channel.

If the channel impulse response is assumed to be time invariant, or is at least stationary in a wide sense, then the baseband channel impulse response can be simplified to

$$h_k(t) = \sum_{p=0}^{P-1} A_{k,p} \exp(-j\theta_{k,p})\delta(\tau - \tau_p). \tag{2.58}$$

It is evident that the discrete time channel impulse response is the summation of a series of impulses, each with a different phase. Applying this to the CDMA case we have the sum of the complex convolutions of the different user signature waveforms with their respective frequency selective channels.

Since we are considering multiple delayed versions of the received signals, we will use an asynchronous equation (2.31). Considering only the ith bit,

$$y(t)[i] = \sum_{k=1}^{K} \sum_{p=0}^{P-1} A_{k,p} b_k[i] s(\tau - \tau_p) \exp(-j\theta_{k,p}) + \sigma n(t). \tag{2.59}$$

where P is the number of equally spaced multipath components.

2.4.2.2 THE CDMA UPLINK AND DOWNLINK CHANNELS

The above model provides for the uplink situation, where the transmission from each mobile to a central base station has a different impulse response. The downlink situation (base station to mobile) in which the channel impulse response pertaining to all the users are equal, i.e., $h_1(t) = h_2(t) = \ldots = h_K(t)$, leads to the simplification that the contribution of all users can be added, before passing through a single channel $h(t)$.

2.5 SUMMARY

This chapter supplies the mathematical background to thoroughly analyze all the multiuser detection schemes contained in this dissertation. The concept of multiuser interference in terms of cross-correlation coefficients is presented here. Discrete synchronous and asynchronous baseband CDMA signal and channel models are introduced in vector matrix notation. The fading mobile channel is also presented in this chapter. The Rayleigh fading and multipath time dispersion models applied to CDMA channels are also introduced.

CHAPTER THREE

THE MATCHED FILTER RECEIVER AND MULTIUSER DETECTION PERFORMANCE MEASURES

This chapter contains the analysis of the single user matched filter. The single user matched filter is the simplest method to demodulate CDMA signals. Several criteria are given by which to measure the performance of multiuser detection schemes. The analysis contained here is largely based on the approach followed by Verdu in [31].
The analyses done in this chapter are done in the real domain. This is to facilitate a geometric understanding of the CDMA multiuser detection problem. In the following chapters, the complex valued CMDA model will be utilized.

3.1 OPTIMAL DECISION RULES AND SUFFICIENT STATISTIC

3.1.1 DECISION RULES AND DECISION REGIONS

To obtain a comprehensive understanding of CDMA detection, we will have to explore the subject of *hypothesis testing* [34]. A certain observed random quantity has a distribution known to belong to a finite set of distributions, each of which is associated with a possible outcome. By sampling and observing the random quantity, we must make a decision as to which distribution (or possible outcome) the sample belongs to. The set of possible outcomes or distributions is often referred to as *hypotheses* in statistical terms. The analysis of the observation or sample is mapped to a decision by means of a *decision rule*. Data demodulation is a hypothesis testing experiment in which the observed quantity is a noise corrupted version of the transmitted signal. There are as many decisions as different values for the transmitted data. For example, in the basic synchronous K user CDMA channel model (2.1), there are 2^K possible decisions, and the observed quantity is a waveform on the interval $[0, T]$.

To make a decision as to what data was transmitted, we need to partition the observation space into *decision regions*, each of which corresponds to a possible transmitted data symbol or hypothesis. Knowledge concerning the distribution of the information source is called *a-priori* knowledge. Let us first assume an equiprobable information source at the transmitter. Assume that within the whole observation space m optimum or non-optimum decision regions $R_i \quad i = 1, \ldots, m$ exist corresponding to m hypotheses. Each hypothesis is distributed according to a probability density function of a random variable Z:

$$\begin{aligned} H_1 &: Z \sim f_{Z|1} \\ &\vdots \\ H_m &: Z \sim f_{Z|m} \end{aligned}$$

When referring to optimum regions, the regions are so chosen that they minimize the error probability. We write the probability of error P_e for arbitrary decision regions, as

$$\begin{aligned} P_e &= 1 - \frac{1}{m}\sum_{i=1}^{m} P[Z \in R_i|i] \\ &= 1 - \frac{1}{m}\sum_{i=1}^{m} \int_{R_i} f_{Z|i}(z)dz \\ &\geq 1 - \frac{1}{m}\int \max_{j=1,\ldots,m} f_{Z|j}(z)dz \end{aligned} \tag{3.1}$$

where the last integral is over the whole observation space. Inequality (3.1) is a lower bound which corresponds to the optimum error probability. There may exist several optimum solutions for the choice of decision boundaries. This non-uniqueness of optimum decision regions arises because there may exist points in the observation space at which the maximum density is achieved by several densities simultaneously. If these elements are arbitrarily assigned to the maximizing hypotheses with the lowest index, we obtain the following optimal decision rule for equiprobable hypotheses.

Proposition 3.1 *(Optimal decision rule - Equiprobable hypotheses) Consider m equiprobable hypotheses under which an observed random vector Z has the following probability density functions*[1]

$$\begin{aligned} H_1 &: Z \sim f_{Z|1} \\ &\vdots \\ H_m &: Z \sim f_{Z|m}, \end{aligned} \tag{3.2}$$

then the following decision regions minimize the error probability

$$R_i = \{z : f_{Z|i}(z) = \max_{j=1,\ldots,m} f_{Z|j}(z)\} - \bigcup_{j=1}^{i-1} R_j, \quad i = 1, \ldots, m. \tag{3.3}$$

[1]The symbol $\sim$ denotes "is distributed according to"

For the case of non-equiprobable transmitted symbols or hypotheses, the a-priori probabilities are denoted as $P[H_i]$. The *a posteriori* probabilities can be computed using Bayes' rule. Conditioned on a particular realization of z of the observation, the conditional (a-posteriori) probabilities for hypothesis H_i is given by

$$P[H_i|z] = \frac{f_{Z|i}(z)P[H_i]}{\sum_{j=1}^{m} f_{Z|j}(z)P[H_j]}. \tag{3.4}$$

In general, the minimum error probability decision rule is termed the *Maximum a posteriori* (MAP) rule, which selects the hypothesis with the highest $P[H_i|z]$. In the case of unknown a-priori probabilities or equiprobable hypotheses (as in (3.3)), the decisions are known as *Maximum Likelihood* (ML) decisions.

Consider the case of a m-hypothesis testing problem where the observation is a Gaussian vector with dimension L, with independent components, and variance equal to σ^2. The distributions under each of the hypotheses are distinguished by their means. For example, the mean of the jth vector component under hypothesis H_i is denoted by a_{ij}. The probability density function corresponding to H_i is given by

$$f_{Z|i}(z_1, \ldots, z_L) = \frac{1}{(2\pi)^{L/2}\sigma^L} \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{L}(z_j - a_{ij})^2\right), \tag{3.5}$$

and the optimum decision regions for equiprobable hypotheses are

$$R_i = \{(z_1, \ldots, z_L) : \sum_{j=1}^{L}(z_j - a_{ij})^2 = \min_{k=1,\ldots,m}\sum_{j=1}^{L}(z_j - a_{kj})^2\} - \bigcup_{j=1}^{i-1} R_j, \tag{3.6}$$

which means that we select the hypothesis whose mean vector is closest to the observed vector in Euclidian distance.

3.1.2 CONTINUOUS-TIME GAUSSIAN SIGNALS

In many hypothesis testing problems, the observed quantity is not a vector as in equations (3.5) and (3.6), but a real valued function over a finite time interval. This is the case with both the synchronous and asynchronous CDMA receivers. Sometimes a structure can be placed at the receiver input so that the decisions are based on functions of the received waveforms (called *observables* or *decision statistics*) which can be either scalars or vectors. In the case of a real valued observed quantity, we invoke the following counterpart to proposition 3.1.

Proposition 3.2 *(Optimal decision rule - Equiprobable hypotheses and real valued functions) Let* $x_1, \ldots, x_m$ *be finite energy deterministic functions defined on an interval* $\mathcal{R}$ *of the real line. Let* $n(t)$ *be white Gaussian noise with unit power spectral density. Consider* m *equiprobable hypotheses.*

$$\begin{aligned} H_1 &: y(t) = x_1(t) + \sigma n(t),\ t \in \mathcal{R} \\ &\vdots \\ H_m &: y(t) = x_m(t) + \sigma n(t),\ t \in \mathcal{R}, \end{aligned} \tag{3.7}$$

then the following decision regions minimize the error probability

$$R_i = \{y = \{y(t),\ t \in \mathcal{R}\} : f[y|x_i] = \max_{j=1,\ldots,m} f[y|x_j]\} - \bigcup_{j=1}^{i-1} R_j, \tag{3.8}$$

where

$$f[y|x_i] = \exp\left(-\frac{1}{2\sigma^2}\int_{\mathcal{R}} [y(t) - x_i(t)]^2 dt\right). \tag{3.9}$$

The function $f[y|x_i]$ in (3.9) is termed the *likelihood function*, and corresponds to the unnormalized conditional probability density function $f_{Z|j}(z)$ in proposition 3.1. As with (3.6), minimizing $[y(t) - x_i(t)]^2$, maximizes (3.6), giving us the *minimum distance* decision region

$$R_i = \{y = \{y(t),\ t \in \mathcal{R}\} : [y(t) - x_i(t)]^2 = \min_{k=1,\ldots,m} [y(t) - x_k(t)]^2\} - \bigcup_{j=1}^{i-1} R_j. \tag{3.10}$$

This means that the decision regions that minimize the error probability are minimum distance regions. The waveform $x_i(t)$ that is closest to y(t) in mean-square distance is inferred.

3.1.3 SUFFICIENT STATISTIC

A function of an observable random variable $Y = g(y)$, which does not depend on any unknown parameters, is called a *statistic*. A *sufficient statistic* can formally and generally be defined as follows [35]. In a statistical inference problem where a parameter Θ is to be inferred[2] on the basis of observations y, we say that a function of the observation $Y = g(y)$ is a sufficient statistic for Θ if the conditional distribution of y given $g(y)$, denoted as $f_{y|Y}$, does not depend on Θ. This means that if Y is observed, then additional information cannot be obtained from y if the conditional distribution of y given Y is free of Θ. We will later see that in the case of a single user receiver, the decision statistic Y is given by

$$Y = \langle y, x_i \rangle = \int_{\mathcal{R}} y(t)x_i(t)dt; \quad i = 1, \ldots, m. \tag{3.11}$$

[2]In hypothesis testing Θ takes a finite or countably infinite number of values, whereas in *estimation* problems, it takes an uncountable number of values

To prove that (3.11) is a sufficient statistic for $\Theta = \{H_1, \ldots, H_m\} \equiv \{x_1, \ldots, x_m\}$, we will need another definition of sufficient statistic termed the *factorization criterion* [35]. If y has a probability density function $f[y; x_i]$, then Y is a sufficient statistic for Θ if and only if

$$f[y; x_i] = g(Y; x_i)h(y), \tag{3.12}$$

where $g(Y; x_i)$ does not depend on y, except through Y, and $h(y)$ does not involve x_i. The proof that Y as defined in (3.11) is a sufficient statistic for Θ, is given by

$$\begin{aligned} f[y; x_i] = f[y|x_i] &= \exp\left(\frac{-1}{2\sigma^2}\int_{\mathcal{R}}[y(t) - x_i(t)]^2 dt\right) \\ &= \exp\left(\frac{-1}{2\sigma^2}\left[\int_{\mathcal{R}} y(t)^2 dt - 2\int_{\mathcal{R}} y(t)x_i(t)dt + \int_{\mathcal{R}} x_i(t)^2 dt\right]\right) \\ &= \exp\left(\frac{-1}{2\sigma^2}\left[\int_{\mathcal{R}} y(t)^2 dt - 2Y + \int_{\mathcal{R}} x_i(t)^2 dt\right]\right) \\ &= \exp\left(\frac{-1}{2\sigma^2}\left[2Y + \int_{\mathcal{R}} x_i(t)^2 dt\right]\right)\exp\left(\frac{-1}{2\sigma^2}\int_{\mathcal{R}} y(t)^2 dt\right) \\ &= g(Y; x_i)h(y), \end{aligned} \tag{3.13}$$

where we have split the function $f[y; x_i]$ into the factors $g(Y; x_i)h(y)$. This satisfies the factorization criterion, and proves that (3.11) contains all the information in the original observations to make an optimal decision.

3.2 THE OPTIMAL RECEIVER - SINGLE USER

In this section we will study the optimal receiver for the single user CDMA channel. For a single user, the channel simplifies to

$$y(t) = Abs(t) + \sigma n(t), \quad t \in [0, T] \tag{3.14}$$

where $s(t)$ is deterministic and has unit energy, the noise term $n(t)$ is white and Gaussian and bit $b \in \{\pm 1\}$. The amplitude of the single user is denoted by A.

3.2.1 LINEAR DETECTORS

Before deriving the optimum demodulator for the single user channel, it is insightful to consider the class of detectors termed *linear detectors*. A detector that outputs the sign of the correlation of the received signal with a deterministic signal $\varphi(t)$ of duration T is given by

$$\hat{b} = \text{sgn}\left(\langle y, \varphi\rangle\right) = \text{sgn}\left(\int_0^T y(t)\varphi(t)dt\right) \tag{3.15}$$

The detector extracts the information contained in the observed waveform $y(t)$ by means of the scalar decision statistic $\langle y, \varphi \rangle$. The decision statistic is given by

$$Y = \langle y, \varphi \rangle = Ab \langle s, \varphi \rangle + \sigma \langle n, \varphi \rangle \tag{3.16}$$

The linearity of the decision statistic makes it easy to discern the respective contributions of signal and noise, whereby the choice of φ can be optimized. Having the signal and noise terms separated, we will attempt to determine an optimum value for φ. A sensible way to do this, is to maximize the signal-to-noise ratio (SNR) γ of the decision statistic Y. The signal variance is simply $A^2(\langle s, \varphi \rangle)^2$. A property of white Gaussian noise is that $E[\langle n, \varphi \rangle^2] = ||\varphi||^2$. The noise variance is thus equal to $\sigma^2||\varphi||^2$. The SNR of the decision statistic Y maximized with respect to φ is given by

$$\gamma_{\max} = \max_{\varphi} \frac{A^2 \left(\langle s, \varphi \rangle \right)^2}{\sigma^2 \left\| \varphi \right\|^2}. \tag{3.17}$$

Equation (3.17) can readily be solved by means of the Cauchy-Schwarz inequality $(\langle s, \varphi \rangle)^2 \leq ||\varphi||^2 ||s||^2$, where the equality is only satisfied if and only if φ is a nonzero multiple α of s. Thus we have the maximized SNR given by

$$\gamma_{\max} = \frac{A^2 \left\| \alpha s \right\|^2 \left\| s \right\|^2}{\sigma^2 \left\| \alpha s \right\|^2} = \frac{A^2 \left\| s \right\|^2}{\sigma^2} \tag{3.18}$$

We conclude that any nonzero multiple α of the signal s will maximize the SNR of the decision statistic Y. This excludes the negative multiples of s, as they will yield erroneous decisions in the absence of noise. The value of the constant will have no effect on the maximum SNR, as well as the decisions

$$\hat{b} = \operatorname{sgn} \left(\langle y, \alpha s \rangle \right) = \operatorname{sgn} \left(\int_0^T y(t) s(t) dt \right). \tag{3.19}$$

The detector in (3.19) is known as the *matched filter* or *conventional* detector. We have seen that the matched filter detector is optimal, in that it maximizes the SNR of the decision statistic Y. A linear filter with an impulse response $s(T - t)$ sampled at multiples of time T is equivalent to the decision statistic $\langle y, \varphi \rangle$ in (3.16).

3.2.2 ERROR PROBABILITY - OPTIMAL SINGLE USER LINEAR DETECTOR

Let us investigate the conditional distributions of the decision statistic Y for a DS-CDMA system with binary antipodal modulation. We assume that the noise term $n(t)$ is a Gaussian process. A property of a Gaussian process $n(t)$, is that the inner product $\langle n, \varphi \rangle$ is a Gaussian random variable. Therefore from (3.16), the decision statistic conditioned on $\{-1, +1\}$ is Gaussian with mean $\{-A\langle y, \varphi \rangle, +A\langle y, \varphi \rangle\}$ respectively. The variance for both distributions is equal to $\sigma^2||\varphi||^2$. The Gaussian conditional distributions of Y is abbreviated by $\mathcal{N}(-A\langle y, \varphi \rangle, \sigma^2||\varphi||^2)$ for a minus one and

$\mathcal{N}(+A\langle y,\varphi\rangle, \sigma^2||\varphi||^2)$ for a one sent, respectively. Figure 3.1 shows the conditional distributions of Y conditioned on the transmission of $b=-1$ and $b=1$.

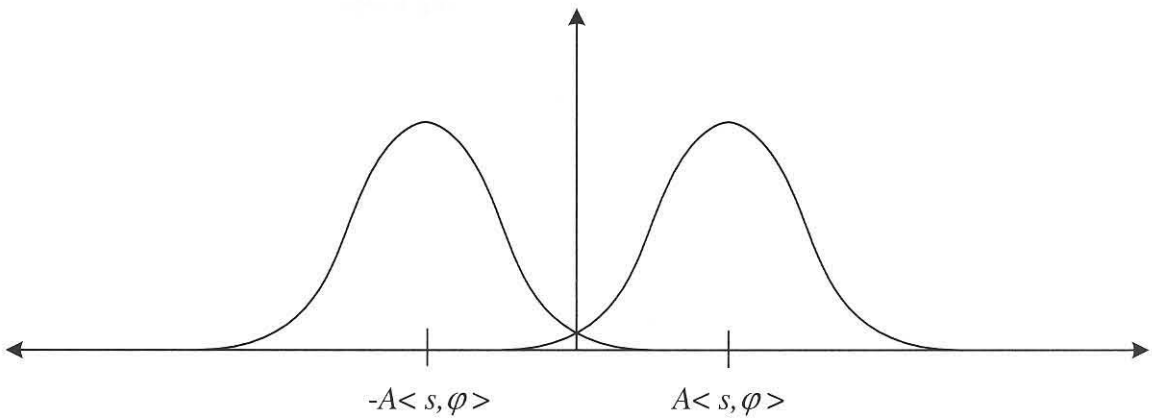

Figure 3.1: Conditional distributions of Y given $b=-1$ and $b=+1$

In the single user binary antipodal case, we have the hypothesis testing problem:

$$\begin{aligned} H_{-1} &: f_{Y|-1} = \mathcal{N}\left(-A\left\langle s,\varphi\right\rangle, \sigma^2||\varphi||^2\right) \\ H_{+1} &: f_{Y|+1} = \mathcal{N}\left(+A\left\langle s,\varphi\right\rangle, \sigma^2||\varphi||^2\right) \end{aligned} \tag{3.20}$$

which is a special case of the vector Gaussian problem in (3.6) with $L=1$ and $m=2$. The corresponding decision regions are

$$R_{-1} = \{y\in(-\infty,\infty): f_{Y|-1}(y) > f_{Y|+1}(y)\} = (-\infty,0) \tag{3.21}$$

$$R_{+1} = \{y\in(-\infty,\infty): f_{Y|+1}(y) > f_{Y|-1}(y)\} = [0,\infty), \tag{3.22}$$

which means that the boundary or threshold between the two regions is at $x=0$. Using the decision regions in (3.21) and (3.22) the probability of error is given by

$$\begin{aligned} P_e &= \frac{1}{2}\int_0^\infty f_{Y|-1}(v)dv + \frac{1}{2}\int_{-\infty}^0 f_{Y|+1}(v)dv \\ &= \frac{1}{2}\int_{A\langle s,\varphi\rangle}^\infty \frac{1}{\sqrt{2\pi}\sigma\left\|\varphi\right\|^2}\exp\left(-\frac{v^2}{2\sigma^2\left\|\varphi\right\|^2}\right)dv \\ &\quad + \frac{1}{2}\int_{-\infty}^{-A\langle s,\varphi\rangle} \frac{1}{\sqrt{2\pi}\sigma\left\|\varphi\right\|^2}\exp\left(-\frac{v^2}{2\sigma^2\left\|\varphi\right\|^2}\right)dv \\ &= \int_{\frac{A\langle s,\varphi\rangle}{\sigma\|\varphi\|}}^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}dv \qquad (3.23) \\ &= Q\left(\frac{A\left\langle s,\varphi\right\rangle}{\sigma\left\|\varphi\right\|}\right), \qquad (3.24) \end{aligned}$$

where (3.23) follows by symmetry and a change of integration variable, and (3.24) follows from the notation of the complementary cumulative distribution function of the unit normal random variable or Q function. Assuming a matched filter receiver, the error probability simplifies to

$$P_e = Q\left(\frac{A}{\sigma}\right) = Q\left(\sqrt{\frac{A^2}{\sigma^2}}\right) = Q\left(\sqrt{\gamma}\right), \tag{3.25}$$

where γ denotes the SNR.

In much of the literature on digital communication systems, bit error probability (BEP) is given in terms of bit energy E_b and N_0 where N_0 is related to noise variance by $\sigma^2 = N_0/2$. The bit energy E_b is simply equal to A^2, since the signature waveform is assumed to have unit energy. Thus the matched filter probability of error can also be written as

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{3.26}$$

which is equal to the BEP of a BPSK system [27].

3.2.3 ERROR PROBABILITY - OPTIMAL SINGLE USER NON-LINEAR DETECTOR

Let us now search for the detector that achieves the minimum error probability among all detectors, by dropping the linearity constraint as imposed in (3.16). This means that we can no longer assume that the observable is $\langle s, \varphi \rangle$ and we have to work with the received process $\{y(t), t \in [0, T]\}$ itself. This is a special case of the problem solved in Proposition 3.2 with $m = 2$, $\mathcal{R} = [0, T]$ and $x_1(t) = As(t)$, $x_2(t) = -As(t)$. Because the energies of x_1 and x_2 are identical, the minimum error probability detector decides $\hat{b} = 1$ if

$$\int_{\mathcal{R}} y(t)x_1(t)dt \geq \int_{\mathcal{R}} y(t)x_2(t)dt, \tag{3.27}$$

and

$$\int_{\mathcal{R}} y(t)x_1(t)dt = -\int_{\mathcal{R}} y(t)x_2(t)dt = A\int_0^T y(t)s(t)dt, \tag{3.28}$$

which means the matched filter output statistic $\langle s, \varphi \rangle$ is a sufficient statistic, and the detector in (3.19) is optimal among all detectors. The shape of the transmitted signal does not affect the minimum bit-error-rate, because of the inherent symmetry of white Gaussian noise, i.e., its projections along every direction has the same distribution.

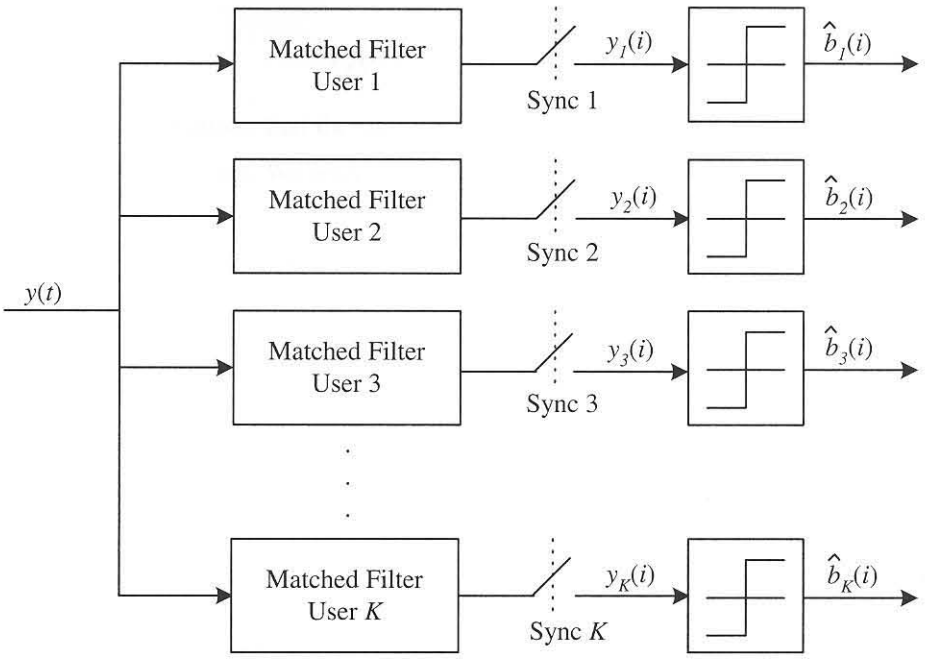

Figure 3.2: Block diagram depicting the bank of matched filters for multiple CDMA users

3.3 MATCHED FILTER ERROR PROBABILITY - SYNCHRONOUS USERS

In this section we will analyze the performance of the single user matched filter in a multiuser CDMA environment. In the multiuser case, demodulation is achieved by a bank of matched filters (Figure 3.2), each matched to a specific user's signature waveform. In the synchronous case we need only to concern ourselves with the timing of a single synchronizer to sample the matched filter outputs of all the users. The output of the kth matched filter in a K user channel is given by

$$y_k = \int_0^T y(t)s_k(t)dt = A_k b_k + \sum_{j \neq k} A_j b_j \rho_{jk} + n_k \tag{3.29}$$

as in equation (2.10), with

$$n_k = \sigma \int_0^T n(t)s_k(t)dt \tag{3.30}$$

a Gaussian random variable with zero mean and variance equal to σ^2. Consider the case of orthogonal signature waveforms, then $\rho_{jk} = 0$ for $j \neq k$, and the problem reduces to the single user case with $y_k = A_k b_k + n_k$. The error probability with orthogonal signature waveforms also reduces to the single user case with

$$P_e(\sigma, k) = Q\left(\frac{A_k}{\sigma}\right), \tag{3.31}$$

which leads us to the conclusion that the matched filter is optimal in a K user CDMA channel with orthogonal signature waveforms. We return to the non-orthogonal CDMA channel.

3.3.1 THE TWO USER CASE

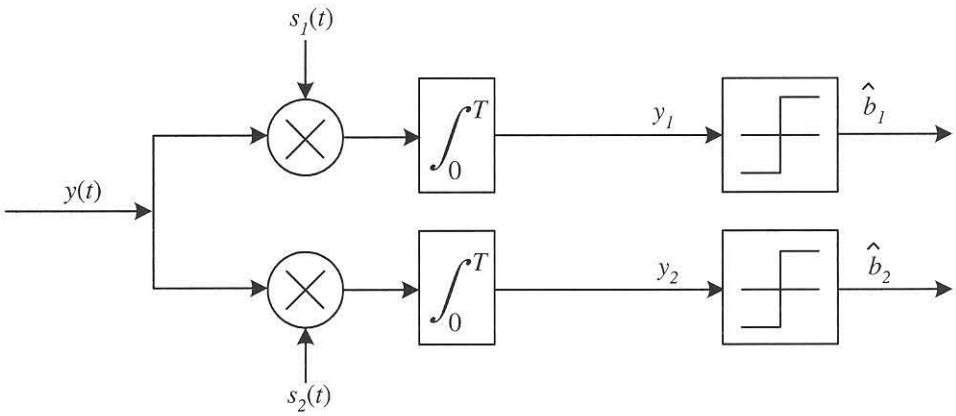

Figure 3.3: Block diagram depicting the special case of the two user CDMA matched filter receiver structure

The two user CDMA channel (Figure 3.3) is instrumental to developing a thorough intuitive and visual understanding of the multiuser interference problem. We start by determining the error probability of user 1 as given by

$$\begin{aligned} P_e(\sigma, 1) &= P\left[\hat{b}_1 \neq b_1\right] \\ &= P\left[b_1 = +1\right] P\left[y_1 < 0 | b_1 = +1\right] \\ &\quad + P\left[b_1 = -1\right] P\left[y_1 \geq 0 | b_1 = -1\right], \end{aligned} \tag{3.32}$$

but y_1 conditioned on b_1 is not Gaussian, so we will have to condition on b_2 as well, with

$$\begin{aligned} P\left[y_1 \geq 0 | b_1 = -1\right] &= P\left[y_1 \geq 0 | b_1 = -1, b_2 = +1\right] P\left[b_2 = +1\right] \\ &\quad + P\left[y_1 \geq 0 | b_1 = -1, b_2 = -1\right] P\left[b_2 = -1\right]. \end{aligned} \tag{3.33}$$

Substitute (3.29) into (3.33) for $\{b_1 = -1, b_2 = +1\}$ and $\{b_1 = -1, b_2 = -1\}$, to obtain

$$\begin{aligned} P[y_1 \geq 0 | b_1 = -1] &= P[n_1 \geq A_1 - A_2\rho]\, P[b_2 = +1] \\ &\quad + P[n_1 \geq A_1 + A_2\rho]\, P[b_2 = -1] \\ &= \frac{1}{2} Q\left(\frac{A_1 - A_2\rho}{\sigma}\right) + \frac{1}{2} Q\left(\frac{A_1 + A_2\rho}{\sigma}\right) \end{aligned} \tag{3.34}$$

where in the two user case, $\rho_{12} = \rho$. Due to the fact that we assumed equiprobable bitstreams b_1 and b_2, and due to symmetry, we get exactly the same expression for $P[y_1 < 0 | b_1 = +1]$. The bit error probability BEP of the conventional receiver with one interfering user is given by

$$\begin{aligned} P_e(\sigma, 1) = P_e(\sigma, 2) &= \frac{1}{2} Q\left(\frac{A_1 - A_2\rho}{\sigma}\right) + \frac{1}{2} Q\left(\frac{A_1 + A_2\rho}{\sigma}\right) \\ &= \frac{1}{2} Q\left(\frac{A_1 - A_2|\rho|}{\sigma}\right) + \frac{1}{2} Q\left(\frac{A_1 + A_2|\rho|}{\sigma}\right) \end{aligned} \tag{3.35}$$

due to the fact that user 1 is arbitrary. Since the Q function is monotonically decreasing, we readily obtain the upper bound

$$P_e(\sigma, 1) \leq Q\left(\frac{A_1 - A_2|\rho|}{\sigma}\right). \tag{3.36}$$

This bound is smaller than $1/2$, provided that the interferer is not dominant, i.e.

$$\frac{A_2}{A_1} < \frac{1}{|\rho|}. \tag{3.37}$$

In this case, because of the asymptotic behavior ($\sigma \to 0$) of the Q function, equation (3.35) is dominated by the term with the smallest argument. Thus, the upper bound (3.36) is an excellent approximation (modulo a factor two) to $P_e(\sigma, 1)$ for all but low SNRs. This implies that the BEP of the conventional receiver behaves like the BEP of a single user system with a reduced SNR, i.e.

$$\gamma_{\text{equiv}} = \left(\frac{A_1 - A_2|\rho|}{\sigma}\right)^2 \tag{3.38}$$

On the other hand, if the relative amplitude of the interferer is such that

$$\frac{A_2}{A_1} > \frac{1}{|\rho|}, \tag{3.39}$$

then the conventional receiver exhibits a highly anomalous behavior called the *near-far problem*. For example, the error probability is not monotonic with σ. When we consider the limit $\sigma \to \infty$, we obtain the error probability from (3.35) as

$$\lim_{\sigma \to \infty} P_e(\sigma, 1) = \frac{1}{2}, \tag{3.40}$$

which is what we would expect from any detector. At the other extreme for $\sigma \to 0$, we get

$$\lim_{\sigma \to 0} P_e(\sigma, 1) = \frac{1}{2}, \tag{3.41}$$

because due to (3.39), as $\sigma \to 0$, the polarity of the output of the matched filter for user 1 tends to be governed by the bitstream of user 2, rather than that of user 1. In this case, a little Gaussian noise is better than no noise. With zero noise, it can be seen that the interference shifts the matched filter output to the wrong side of the threshold, as in Figure 3.4. The addition of noise can have one of three effects on the decision,

1. no effect,
2. to prevent an error and
3. to induce an error.

The noise sample amplitude needed for 3. is at least $|\rho|A_2 + A_1$, whereas the noise excursion for 2. is only $|\rho|A_2 - A_1$.

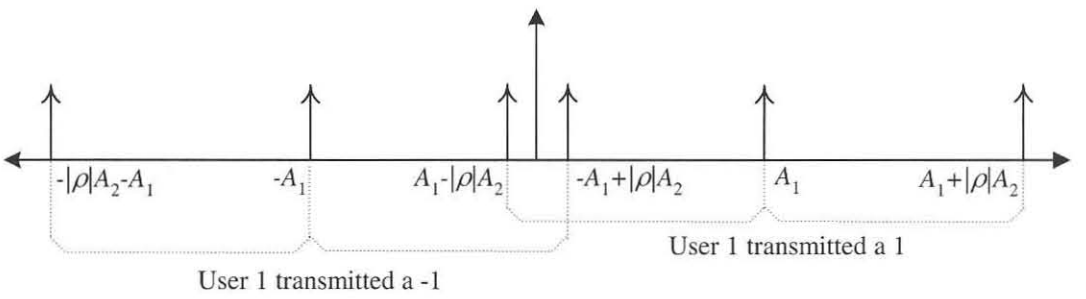

Figure 3.4: Output of the matched filter with one interfering user and $A_2/A_1 > 1/|\rho|$

The noise level that minimizes the BEP under (3.39) is (from [31]):

$$\sigma^2 = \frac{A_1 A_2 \rho}{\operatorname{arctanh}\left(\frac{A_1}{A_2 \rho}\right)} \tag{3.42}$$

Finally we consider the case of equality with

$$\frac{A_2}{A_1} = \frac{1}{\rho}. \tag{3.43}$$

Then the error probability of the single user matched filter reduces to

$$P_e(\sigma, 1) = \frac{1}{4} + \frac{1}{2} Q\left(\frac{2A_1}{\sigma}\right) \tag{3.44}$$

which means that the signal of user 2 exactly cancels the signal of user 1 with a probability of $\frac{1}{2}$ at the matched filter output. It becomes a zero mean Gaussian random variable; with probability $\frac{1}{2}$, the signal of user 2 doubles the contribution of the desired signal to the matched filter output. With respect to the two user case, we will now consider methods of using the BEP as a performance measure. This will give us insight and intuition when considering the K user scenario.

3.3.1.1 BEP AS PERFORMANCE MEASURE - THE TWO USER CASE

When evaluating the performance of digital communication systems, the BEP with respect to the SNR, or alternatively E_b/N_0, is commonly used in the literature. Figure 3.5 shows the BEP for the two user matched filter detector with $\rho = 0.2$ and different relative amplitude values for A_1 and A_2. It can be seen that the BEP degrades rapidly as the relative amplitude of the interferer increases. The top curve is an example of the near-far problem under the condition (3.39).

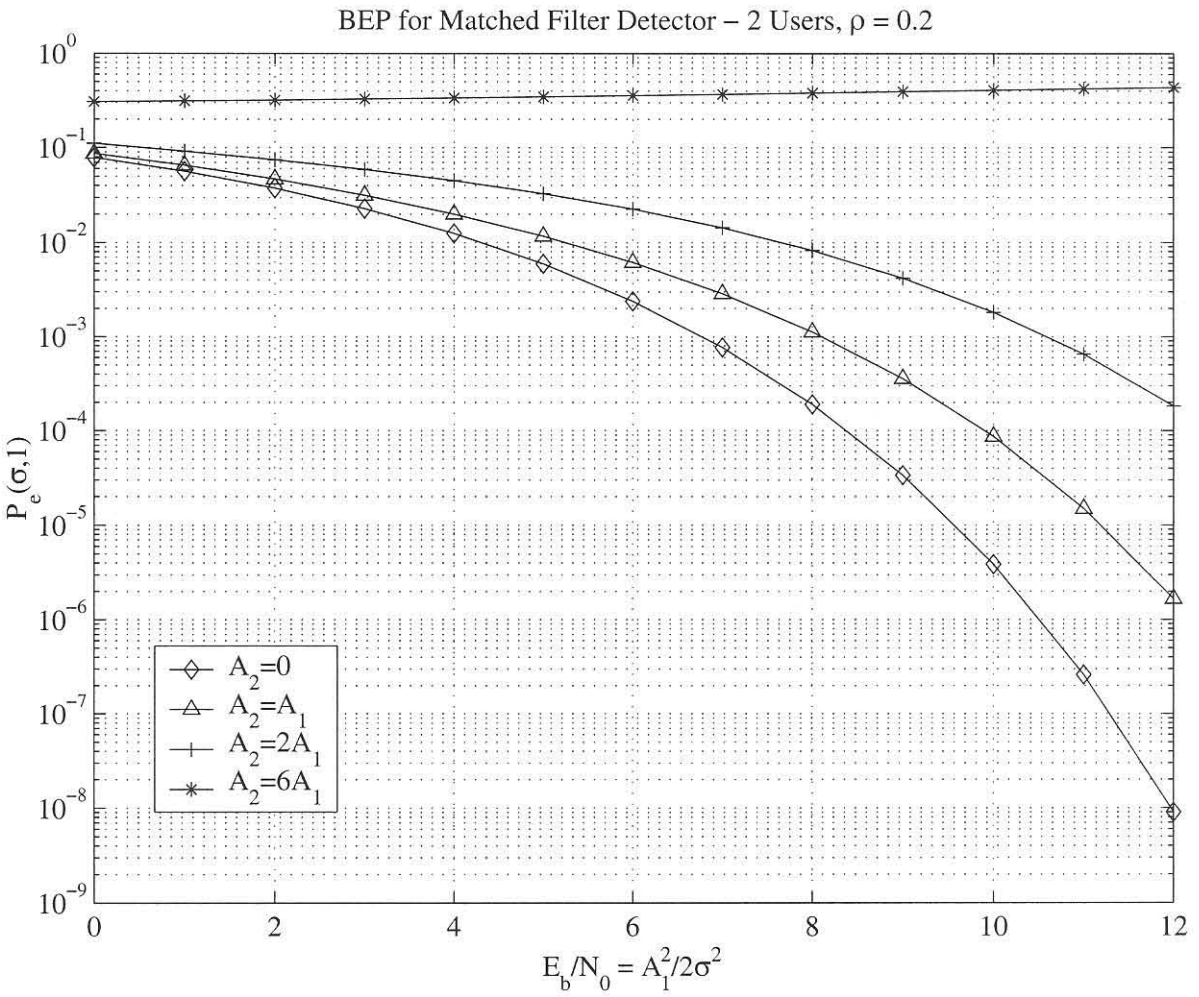

Figure 3.5: BEP of the matched filter detector for different relative amplitudes and $\rho = 0.2$

It is often the case that a digital communication system needs to be designed with a maximum tolerable BEP in mind. The necessary bit energies then need to be found to satisfy that BEP. Figure 3.6 represents the *power-tradeoff* regions so that both users have a BEP of 1×10^{-5}, with the cross-correlation between the two users characterized on the z-axis. In the case of orthogonal users, the objective will be reached for both users if their SNRs are greater than $Q^{-1}(3 \times 10^{-5}) = 12\text{dB}$. From Figure 3.6 it can be seen that as the cross correlation increases:

- even at equal amplitudes the necessary signal energy increases rapidly;
- the sensitivity to imbalances in the received signal grows, making power control necessary.

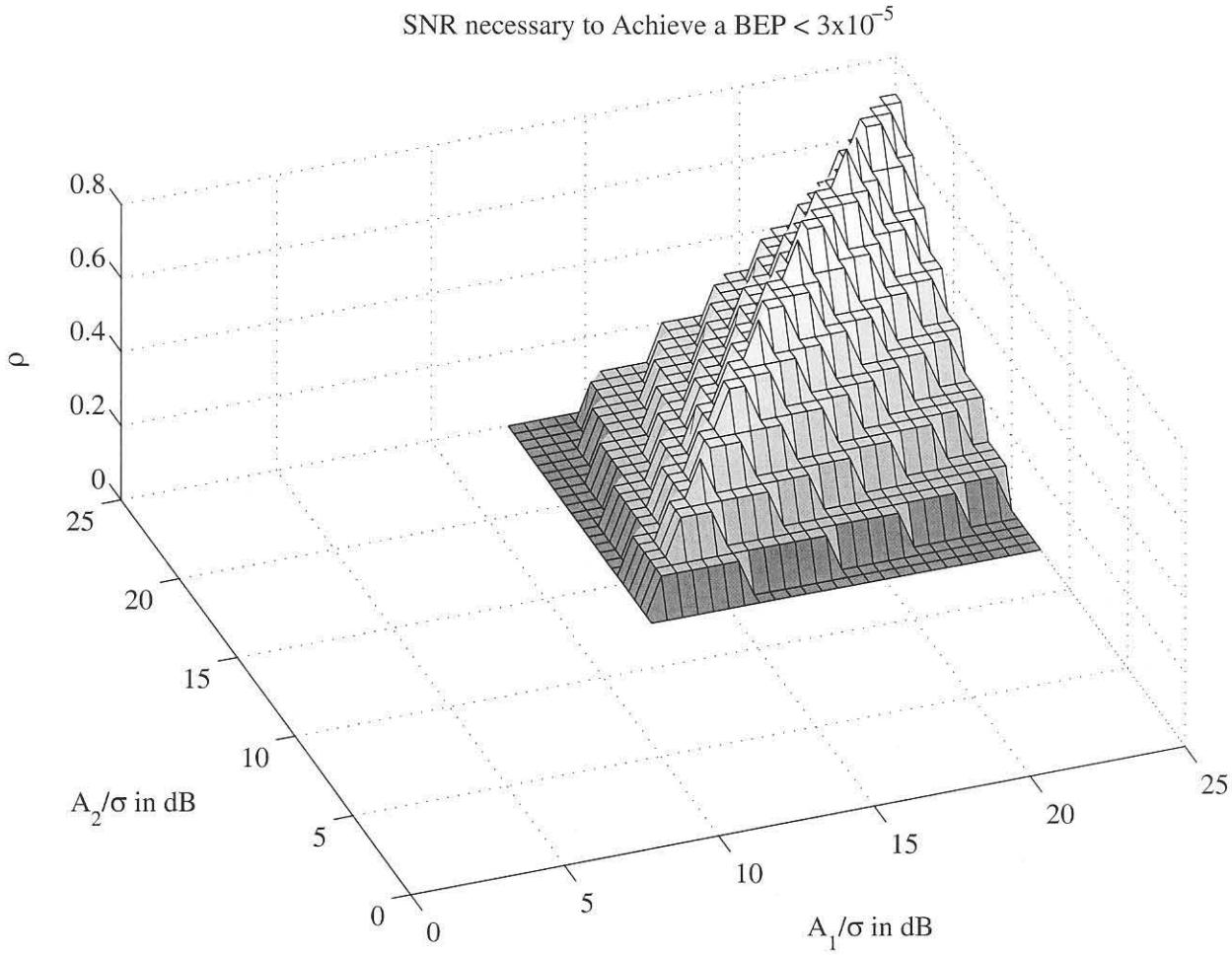

Figure 3.6: Regions of signal-to-noise ratios to attain a BEP of 3×10^{-5} for both users

3.3.1.2 THE TWO USER SIGNAL SPACE REPRESENTATION

Verdu [31] mentions another useful visualization of CDMA detector operation that involves decision regions on a signal space diagram. The signal space representation of detector operation was conceived by Shannon [36] and popularized in the textbook of Wozencraft and Jacobs [37].

For a K user synchronous channel, there are 2^K hypotheses within the observation space on $[0, T]$. This space has infinite dimensions, but the conventional K user demodulator has decision vector space of K dimensions

$$(y_1, \ldots, y_K) = \left(\int_0^T y(t)s_1(t)dt, \ldots, \int_0^T y(t)s_K(t)dt \right). \tag{3.45}$$

To represent the decision regions on a signal space diagram, we will need K dimensions or axes. It is obvious that the two user (two dimensional) case will yield a practical visualization of the signal space. In this case, (y_1, y_2) conditioned on (b_1, b_2) is a Gaussian vector (3.29), (3.30) with mean

$$(A_1 b_1 + A_2 b_2 \rho, A_2 b_2 + A_1 b_1 \rho) \tag{3.46}$$

and covariance matrix

$$\text{cov}(y_1, y_2) = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{3.47}$$

In the (y_1, y_2) signal space (Figure 3.7), we can depict each of the mean vectors for each of the four hypotheses where $A_1 = A_2 = 1$ and $\rho = 0.2$.

The received vector can be viewed as the sum of the transmitted vector (3.46) and a zero mean Gaussian vector (n_1, n_2). The two user received joint Gaussian vector density functions for all four hypotheses is depicted in Figure 3.8 with $\rho = 0.2$ and $\sigma = 1$.

In the absence of noise, as depicted in Figure 3.7, the detector will make correct decisions, since the signal points lie in the correct regions. The probability of error found in (3.34) is the average of the probabilities that the received vector satisfies $y_1 < 0$ given that $(+, +)$ and $(+, -)$ has been transmitted. There is a shortcoming in the (y_1, y_2) signal space diagram in Figure 3.7 in that the noise components (n_1, n_2) are correlated, i.e.

$$E[n_1 n_2] = \sigma^2 \rho \tag{3.48}$$

This has the consequence that the noise vector is not symmetric, nor does the norm of the noise vector determine the likelihood of that realization. This can be seen in Figure 3.9 in the 'overhead' view of Figure 3.8.

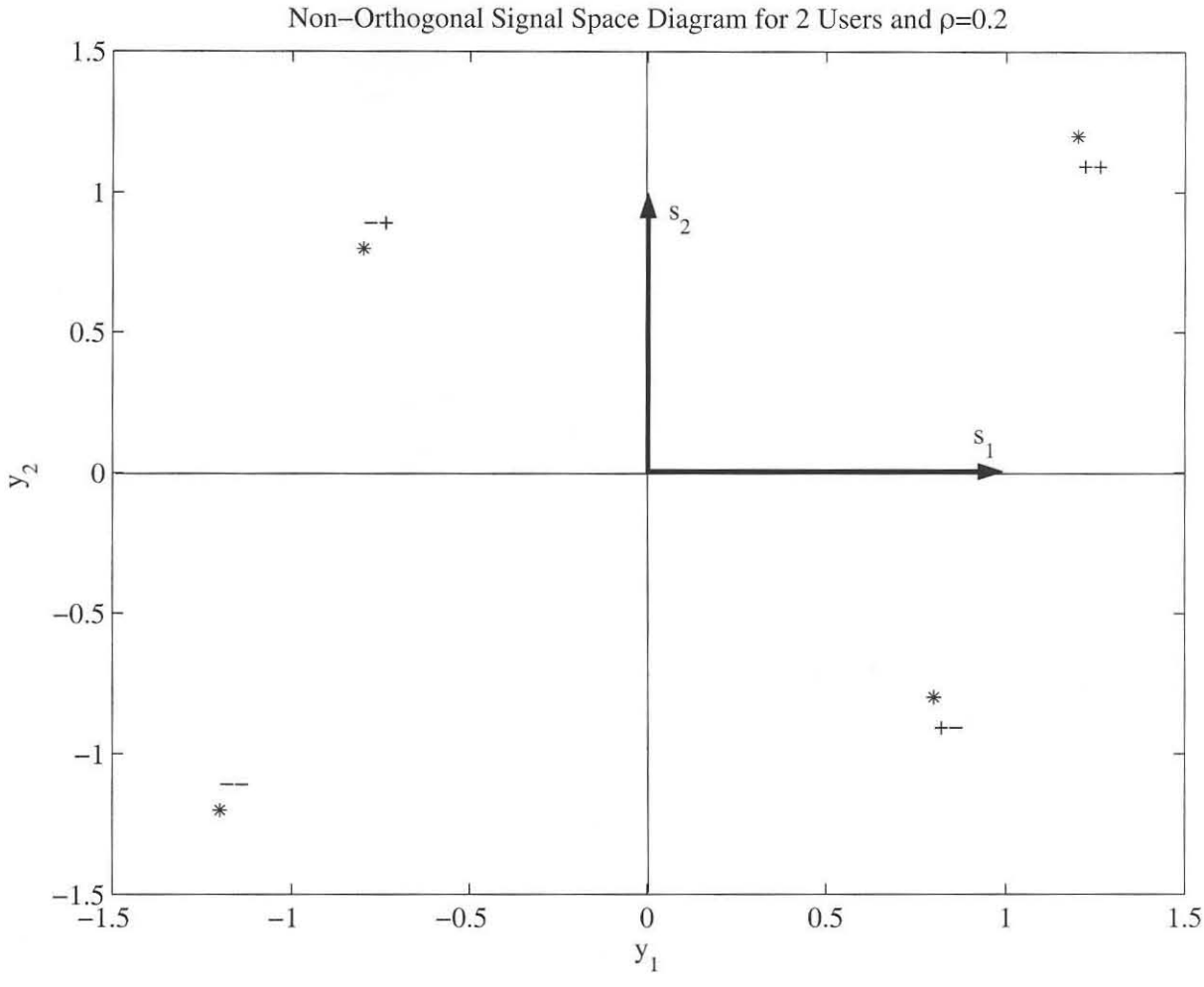

Figure 3.7: Signal space diagram in the (y_1, y_2) space for equal amplitudes and $\rho = 0.2$

A more suitable diagram than the (y_1, y_2) signal space diagram is the $(\tilde{y}_1, \tilde{y}_2)$ signal space diagram whose axes are equal to the correlations of the received waveform with an arbitrary orthonormal basis (ψ_1, ψ_2) that spans the linear space generated by the signals (s_1, s_2). For example, a choice for that orthonormal basis by means of the Gram-Schmidt procedure is

$$\psi_1 = s_1 \tag{3.49}$$

$$\psi_2 = \frac{1}{\sqrt{1-\rho^2}} s_2 - \frac{\rho}{\sqrt{1-\rho^2}} s_1. \tag{3.50}$$

Conditioned on (b_1, b_2), (y_1, y_2) is Gaussian with mean

$$\begin{aligned} &\left(A_1 b_1 \langle s_1, \psi_1 \rangle + A_2 b_2 \langle s_2, \psi_1 \rangle, A_1 b_1 \langle s_1, \psi_2 \rangle + A_2 b_2 \langle s_2, \psi_2 \rangle\right) \\ &= \left(A_1 b_1 + A_2 b_2 \rho, A_2 b_2 \sqrt{1-\rho^2}\right) \end{aligned} \tag{3.51}$$

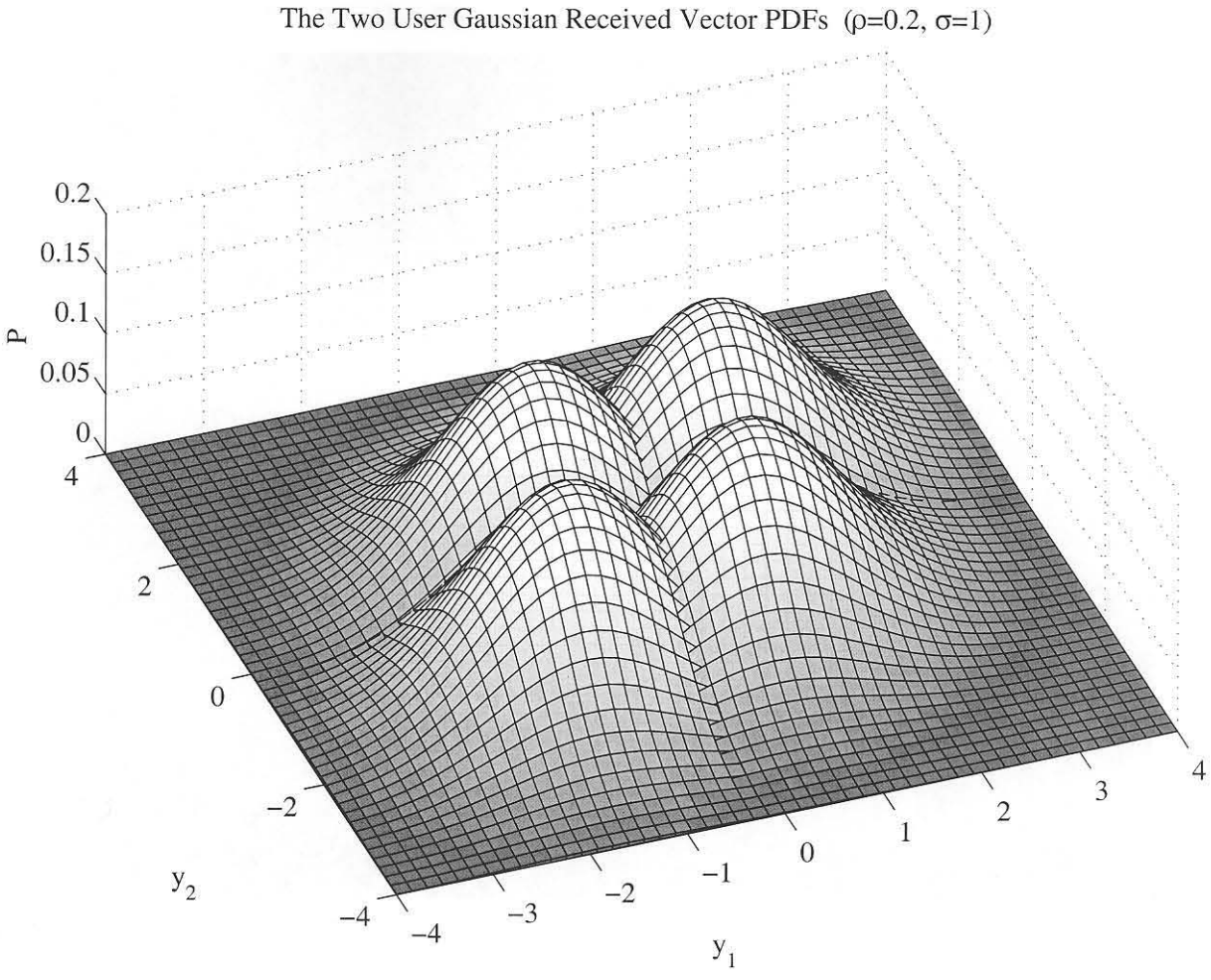

Figure 3.8: Joint probability density function in the (y_1, y_2) space for equal amplitudes, $\rho = 0.2$ and $\sigma = 1$

and covariance matrix equal to

$$\text{cov}(\tilde{y}_1, \tilde{y}_2) = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}. \tag{3.52}$$

The whitened counterpart to Figure 3.7 in alternative orthogonal representation $(\tilde{y}_1, \tilde{y}_2)$ is shown in Figure 3.10. Here, the decision regions are defined by the lines (or hyperplanes in K dimensional space) orthogonal to s_1 and s_2 respectively. With the alternative representation, the inner product between the vectors representing the signature waveforms s_1 and s_2 in Figure 3.10 are, in contrast with Figure 3.7, indeed equal to their cross-correlation.

Even though $(\tilde{y}_1, \tilde{y}_2)$ are not computed by the detector, it is useful to visualize the received vector as belonging to the alternative orthogonal two dimensional space. Indeed, the dimensions of the detector in Figure 3.10 are transparent to all the infinite components in $y(t)$ orthogonal to ψ_1 and ψ_2.

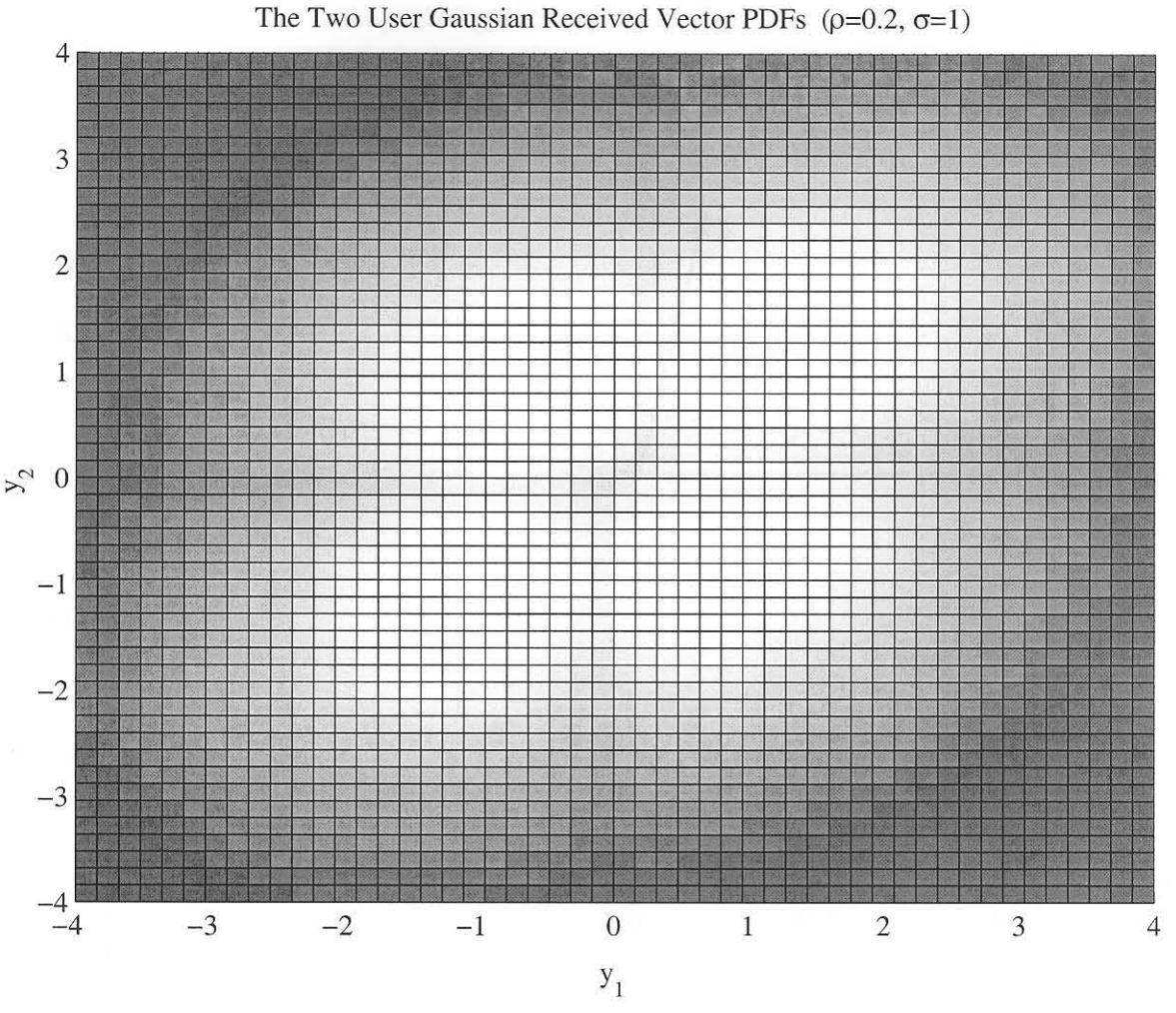

Figure 3.9: Overhead view of the joint probability density function in the (y_1, y_2) space for equal amplitudes, $\rho = 0.2$ and $\sigma = 1$

The anomalous behavior of the conventional matched filter detector in the near-far situation in (3.39) is illustrated in Figure 3.11 with $A_2 = 6A_1$. The decision regions stay exactly the same as in Figure 3.10. The transmitted vectors corresponding to $(+,-)$ and $(-,+)$ have now migrated outside the correct decision regions. This means that given $(+,-)$ or $(-,+)$ was transmitted, an error will occur unless the noise realization moves the vector back into the correct decision region. In a noiseless environment, the decisions of both users is equal to the data transmitted by user 1.

3.3.2 THE K-USER CASE

In the generalization of the BEP to the K user case, we will follow a similar approach a in the two user case. Following the same reasoning as before, the kth user BEP is given by

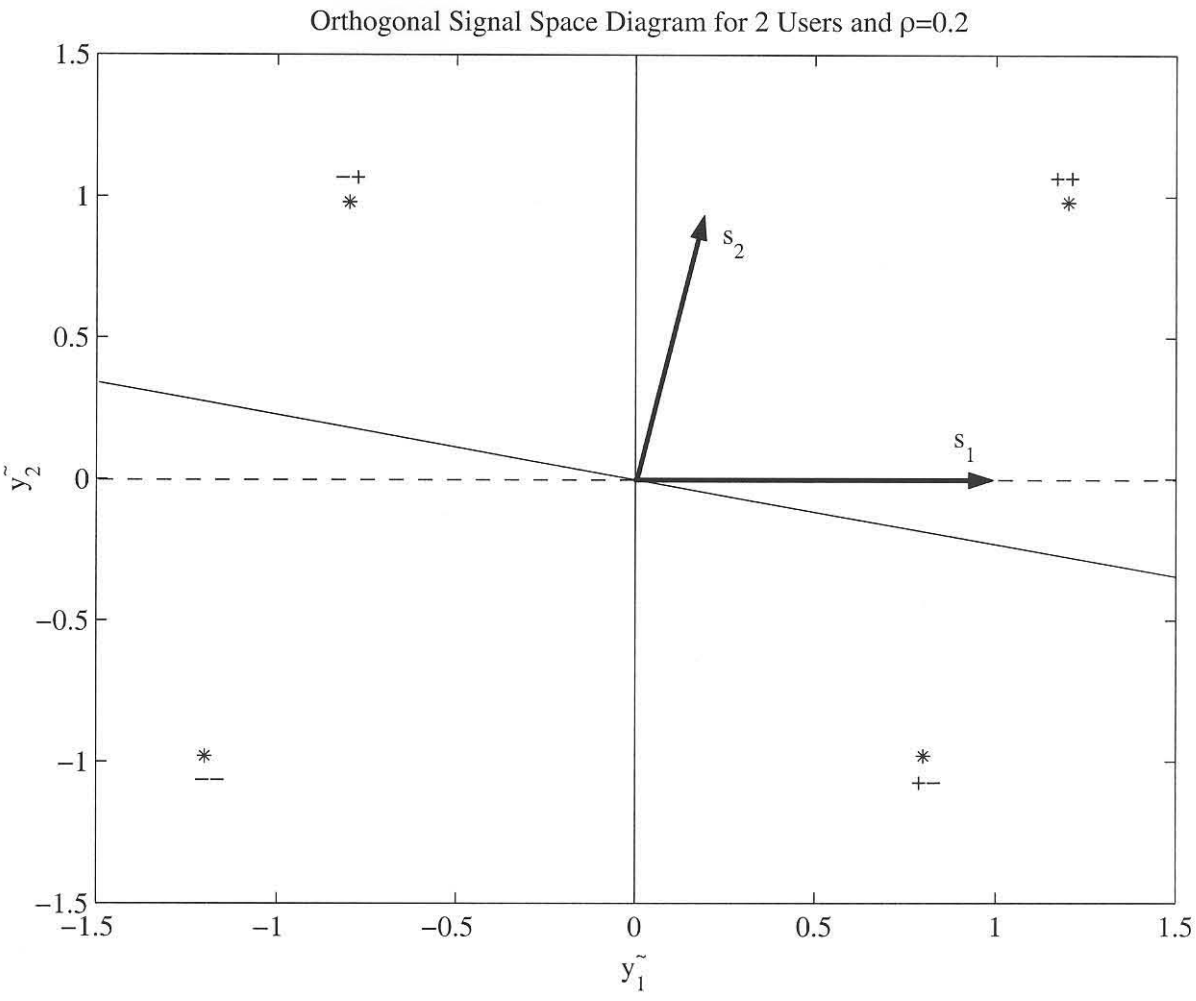

Figure 3.10: Signal space diagram in the alternative orthogonal $(\tilde{y}_1, \tilde{y}_2)$ space for equal amplitudes $A_1 = A_2$ and $\rho = 0.2$

$$\begin{aligned} P_e(\sigma, k) &= P\left[b_k = 1\right] P\left[y_k < 0 | b_k = 1\right] \\ &\quad + P\left[b_k = -1\right] P\left[y_k > 0 | b_k = -1\right] \\ &= \frac{1}{2} P\left[n_k < -A_k - \sum_{j \neq k} A_j b_j \rho_{jk}\right] \\ &\quad + \frac{1}{2} P\left[n_k > A_k - \sum_{j \neq k} A_j b_j \rho_{jk}\right]. \end{aligned} \tag{3.53}$$

Because of the symmetry of the two terms in (3.53), they are equal, and the BEP of the kth user becomes

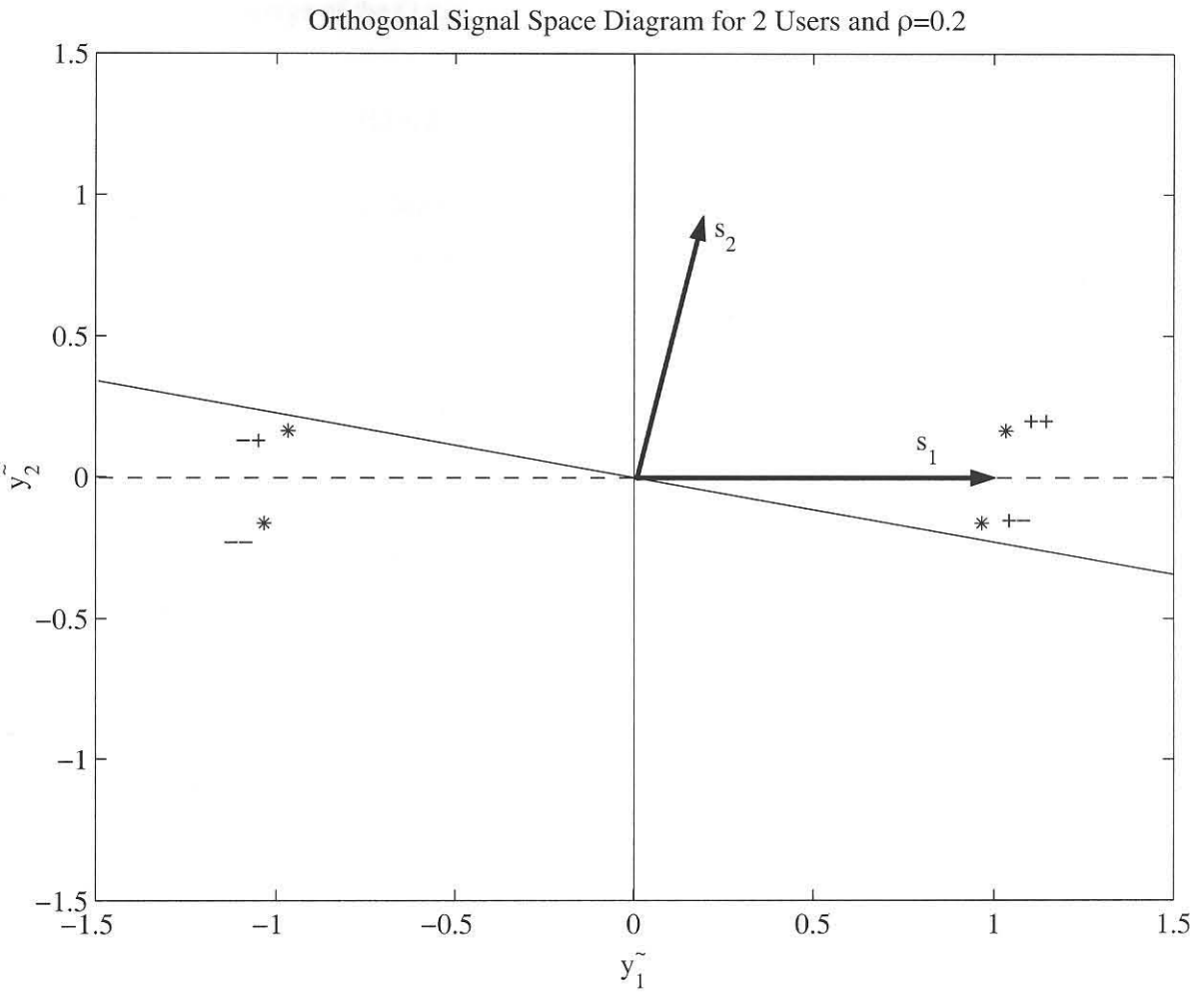

Figure 3.11: Signal space diagram in the alternative orthogonal $(\tilde{y}_1, \tilde{y}_2)$ space for $A_2 = 6A_1$ and $\rho = 0.2$

$$P_e(\sigma, k) = P\left[n_k > A_k - \sum_{j \neq k} A_j b_j \rho_{jk}\right] \tag{3.54}$$

$$= \frac{1}{2^{K-1}} \sum_{(b_1,\ldots,b_K)=(\{-1,1\},\ldots,\{-1,1\})} Q\left(\frac{A_k}{\sigma} + \sum_{j \neq k} b_j \frac{A_j}{\sigma} \rho_{jk}\right) \tag{3.55}$$

where (3.55) is conditioned on all the interfering bits. We see from equation (3.55) that the kth user error probability depends only on the shape of the signature waveforms through their cross-correlations over the interval $[0, T]$, as determined by the receiver. This is also due to the fact that the noise is white and Gaussian. The error probability, as in all digitally modulated systems, depend on the SNR $\frac{A_k}{\sigma}$ and in the CDMA case on the relative amplitudes of the interfering users. As in (3.36),

error probability or average of the Q functions in (3.55) is upper bounded by

$$P_e(\sigma, k) \leq Q\left(\frac{A_k}{\sigma} - \sum_{j \neq k} \frac{A_j}{\sigma} |\rho_{jk}|\right). \tag{3.56}$$

When we look at the anomalous behavior of the condition (3.39) in the K user case, we note that (3.55) goes to zero as $\sigma \to 0$ if and only if the argument of each of the Q functions therein is positive, that is if

$$A_k > \sum_{j \neq k} A_j |\rho_{jk}| \tag{3.57}$$

The condition in (3.57) is commonly referred to as the *open eye* condition. Under this condition, the bound (3.56) becomes tight (modulo a factor independent of σ) as $\sigma \to 0$.

3.3.3 THE GAUSSIAN APPROXIMATION FOR BEP

Equation (3.55) is cumbersome in the sense that the number of operations required increases exponentially with the number of users. It is for this reason that a number of authors, including the classical papers of Pursley [38] and Yao [39], have approximated (3.55) by replacing the binomial random variable

$$\sum_{j \neq k} A_j b_j |\rho_{jk}| \tag{3.58}$$

with a Gaussian random variable with identical variance. The Gaussian approximated BEP becomes

$$\tilde{P}_e(\sigma, k) = Q\left(\frac{A_k}{\sqrt{\sigma^2 + \sum_{j \neq k} A_j^2 \left|\rho_{jk}^2\right|}}\right). \tag{3.59}$$

The approximation in (3.59) is fairly accurate at low SNRs, but for high SNRs it may become more unreliable. A comparison of the exact BEP versus the Gaussian approximation is shown in Figure 3.12 and Figure 3.13 for 10 and 14 equal energy users, respectively. The cross correlation ρ is set at 0.08. Figure 3.12 is representative of the open eye situation and Figure 3.13 of the closed eye situation. In the latter case we notice that the behavior of the BEP of the single user matched filter detector is non-monotonic. This was also observed in the two user case as the "anomalous" near far situation. In the limit as $\sigma \to 0$, equations (3.55) and (3.59) behave differently. Equation (3.59) has a nonzero limit, even if the open eye condition is satisfied. The reason for this is that when we approximate the binomial random variable with a Gaussian random variable, the error is greatest in the tails, which determine the BEP for high SNRs. When the performance is averaged with respect

to random carrier phases, the multiuser interference is no longer binomially distributed, but remains amplitude limited. This is in contrast to a Gaussian random variable with the same variance.

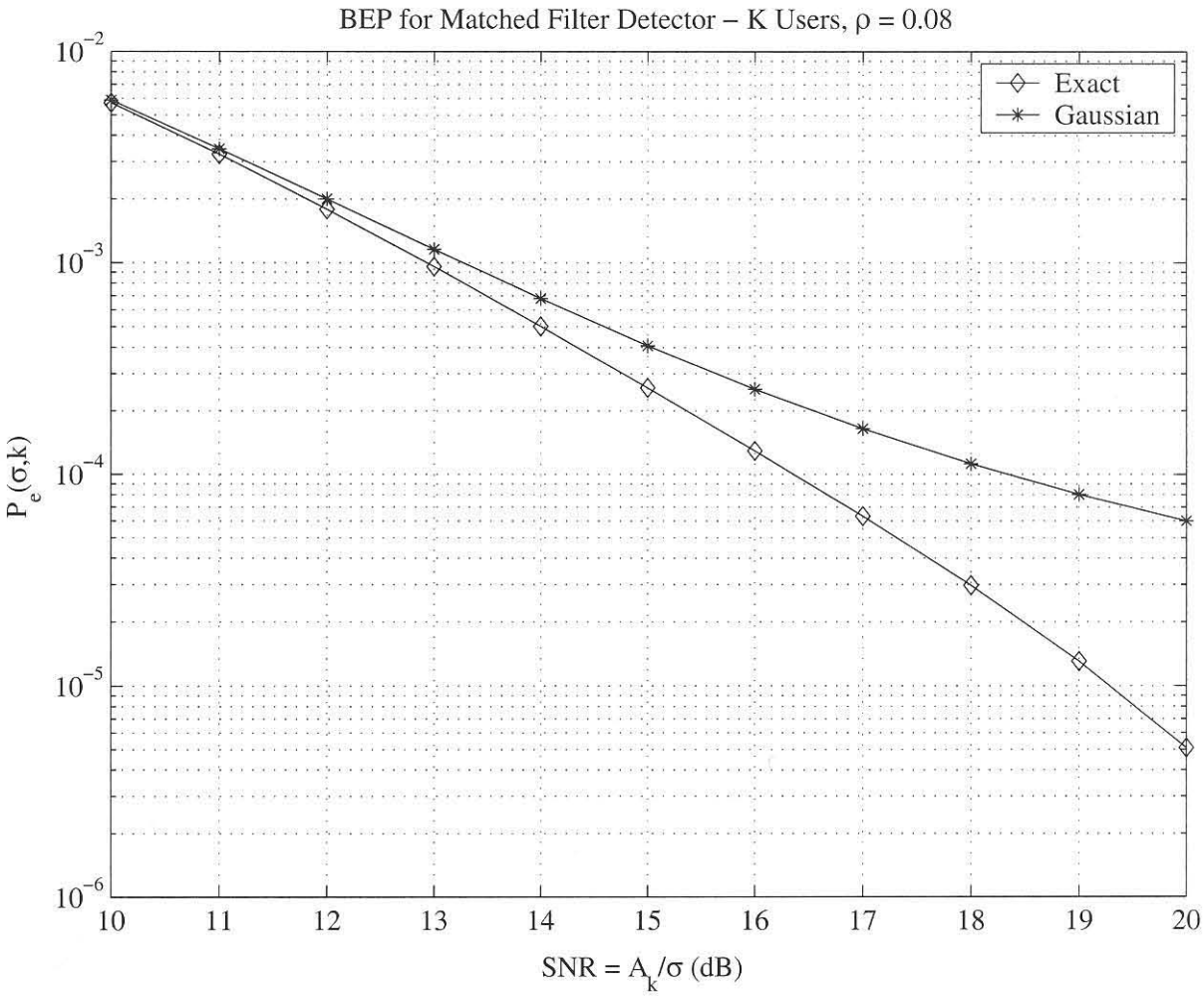

Figure 3.12: BEP as a function of SNR with $K = 10$ equal energy users and $\rho = 0.08$ (eye open)

Proposition 3.3 *Suppose that the random direct sequence model is used and BEP is averaged with respect to the choice of binary sequences with spreading gain N. If $K \to \infty$ and $N \to \infty$, but their ratio is kept constant*

$$\frac{K}{N} = \beta, \tag{3.60}$$

then the averaged BEP converges to

$$\lim_{K,N\to\infty} E\left[P_e\left(\sigma,1\right)\right] = Q\left(\frac{A_1}{\sqrt{\sigma^2+\beta\bar{A}^2}}\right) \tag{3.61}$$

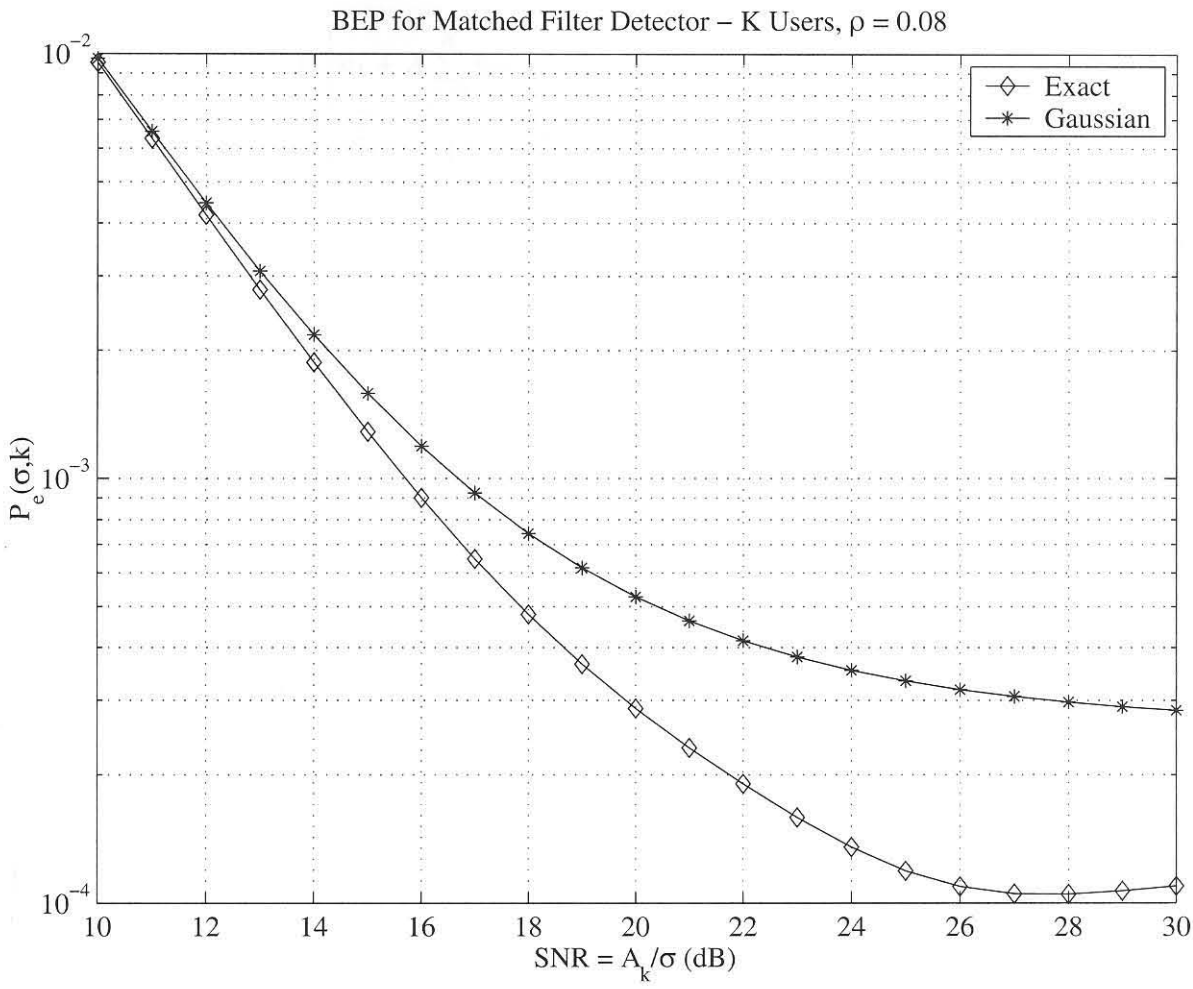

Figure 3.13: BEP as a function of SNR with $K = 14$ equal energy users and $\rho = 0.08$ (eye closed)

where

$$\bar{A}^2 \stackrel{\text{def}}{=} \lim_{K\to\infty} \frac{1}{K} \sum_{j=2}^{K} A_j^2 \tag{3.62}$$

A sufficient condition for the validity of (3.62) is that the amplitudes A_j be bounded.

Let us justify (3.62) under the condition that all energies are equal for all users, i.e. $A_k = A$. According to (3.54) we need to compute the limit of

$$\begin{aligned} & P\left[n_1 + A\sum_{j=2}^{K} b_j\rho_{1j} > A\right] \\ &= P\left[n_1 + A\sum_{j=2}^{K} b_j\left(\tfrac{1}{N}\sum_{n=1}^{N} d_{jn}\right) > A\right] \\ &= P\left[n_1 + A\sqrt{\tfrac{K-1}{N}}\cdot\tfrac{1}{\sqrt{(k-1)N}}\sum_{j=2}^{K}\sum_{n=1}^{N} d_{jn} > A\right] \end{aligned} \tag{3.63}$$

where the random variables d_{jn} in (3.63) are independent and equally likely to be $\{-1, +1\}$. The De Moivre-Laplace Central Limit Theorem dictates convergence in distribution as $K \to \infty$ of the random variable

$$\frac{1}{\sqrt{(k-1)N}}\sum_{j=2}^{K}\sum_{n=1}^{N} d_{jn} \tag{3.64}$$

to a zero mean, unit variance Gaussian random variable. The right side of (3.63) converges to

$$Q\left(\frac{A}{\sqrt{\sigma^2 + \beta A^2}}\right), \tag{3.65}$$

which is what we wanted to verify.

The limiting result in (3.61) can be strengthened to show that even if the BEP is not averaged with respect to random sequences, it converges as $K = \beta N \to \infty$ to the right side of (3.61) with probability one for any signal to noise ratio [40]. It must be said however, that convergence is very slow with K for high SNRs. An easily computable upper bound to $P_e(\sigma, k)$ can be found by partitioning the set of users into

$$\{1, \ldots, K\} = \{k\} \cup G \cup \bar{G} \tag{3.66}$$

where G is a subset of interferers that satisfies the partial eye open condition, i.e.

$$A_k > \sum_{j\in G} A_j\,|\rho_{jk}|. \tag{3.67}$$

Then the error probability of the single user matched filter is bounded by

$$P_e(\sigma, k) \leq \exp\left(-\frac{\left(A_k - \sum_{j\in G} A_j\,|\rho_{jk}|\right)^2}{2\left(\sigma^2 + \sum_{j\in\bar{G}} A_j^2\rho_{jk}^2\right)}\right) \tag{3.68}$$

of which the justification is given in [31]. This bound is known as the *Chernoff bound* [41]. The freedom to choose G subject to (3.67) can be exploited to minimize the upper bound in (3.68). The conditions of $G = \emptyset$ and $\bar{G} = \emptyset$ deserve special attention.

First, if the fully open eye condition in (3.57) is satisfied and $\bar{G} = \emptyset$, then

$$P_e(\sigma, k) \leq \exp\left(-\frac{\left(A_k - \sum_{j \neq k} A_j \left|\rho_{jk}\right|\right)^2}{2\sigma^2}\right). \tag{3.69}$$

Second, we can set $G = \emptyset$, then (3.68) becomes (cf. (3.59))

$$P_e(\sigma, k) \leq \exp\left(-\frac{A_k^2}{2\left(\sigma^2 + \sum_{j \neq k} A_j^2 \rho_{jk}^2\right)}\right). \tag{3.70}$$

Equations (3.69) and (3.70) are the two extreme conditions for the upper bound of the single-user matched filter BEP.

There have been several other attempts to find better approximations for CDMA BEP bounds for random signature sequences. Some of these are presented in [42], [43] and [44].

3.4 MATCHED FILTER ERROR PROBABILITY - ASYNCHRONOUS USERS

In an asynchronous CDMA system where all users use the same basic chip waveform, the continuous-time to discrete-time conversion can be carried out by a single chip matched filter sampled at K times the chip rate, with the sampling instants determined by the synchronizers.

The analysis of the asynchronous case is identical, except for the fact that each bit is affected by $2K - 2$ interfering bits. This doubles the number of terms in (3.55)

$$P_e(\sigma, k) = \frac{1}{4^{K-1}} \sum_{((b_1,d_1),\ldots,(b_K,d_K))=(\{-1,1\}^2,\ldots,\{-1,1\}^2)} Q\left(\frac{A_k}{\sigma} + \sum_{j \neq k} \frac{A_j}{\sigma}(b_j \rho_{jk} + d_j \rho_{kj})\right). \tag{3.71}$$

The condition in (3.57) can be extended to the asynchronous case,

$$A_k > \sum_{j \neq k} A_j (|\rho_{jk}| + |\rho_{kj}|) \tag{3.72}$$

The asynchronous cross correlations in (3.71) depend on the relative timing offset between users. These parameters are time varying random variables. Given a set of signature waveforms, it is possible to compute the distribution (or simply expectation) of (3.71). This however is computationally

intensive.

The infinite user limit as $K \to \infty$ can be extended to the asynchronous case by incorporating two fictitious interferers per actual interferer. Averaging over the received uniformly distributed delays and considering that the autocorrelation for rectangular chip waveforms is

$$R_p(\tau) = 1 - \tfrac{\tau}{T_c}, \quad 0 \leq \tau \leq T_c \tag{3.73}$$

we can obtain the second moment of the asynchronous cross correlations ρ_{jk} and ρ_{kj} [31]

$$\begin{aligned} \frac{1}{T}\int_0^T E\left[\rho_{jk}^2(\tau)\right] dt &= \frac{1}{NT_c}\int_0^{T_c} R_p^2(\tau) d\tau \\ &= \frac{1}{N}\int_0^1 (1-x)^2 dx \\ &= \frac{1}{3N}. \end{aligned} \tag{3.74}$$

The second moment of ρ_{kj} is equal to that of ρ_{jk} due to symmetry and a uniformly distributed delay. This implies that the BEP is equivalent to that of a synchronous system with $(2/3) \times (K-1)$ interferers.

3.5 ASYMPTOTIC MULTIUSER EFFICIENCY AND RELATED MEASURES

We already considered BEP as a performance measure for the multiuser CDMA environment. There are several other performance measures that can be derived from BEP that will be of value in the comprehension of CDMA detector operation. One such performance measure mentioned earlier, is the power tradeoff region of SNRs that results in a given guaranteed BEP level.

When we consider a slowly time varying channel with respect to delays, phases, and most importantly, SNRs, averaging BEPs may be misleading. This is due to the fact that the channel may be dominated by particularly unfavorable, but rare channel conditions. It is common practice to design a digital communication system with *outage* as design parameter. Outage is defined as the percentage of time that the system performs below a certain level. When designing according to outage as design parameter, the cumulative distribution function of the BEP is more informative than its average.

In this section we will consider signal to interference ratio, multiuser efficiency, asymptotic multiuser efficiency and near-far resistance as CDMA multiuser detector performance measures.

interference caused by other users after detection. We can achieve this by letting $\sigma \to 0$ in (3.81). The *asymptotic multiuser efficiency* of user k is defined in [2] and [45] as

$$\eta_k = \lim_{\sigma \to 0} \frac{e_k(\sigma)}{A_k^2} \tag{3.82}$$

and is the log BEP of the kth user going to zero with the same slope as that of a single user with energy $\eta_k A_k^2$. That is,

$$\eta_k = \sup \left\{ 0 \leq r \leq 1 : \lim_{\sigma \to 0} P_e(\sigma, k)/Q\left(\frac{\sqrt{r}A_k}{\sigma}\right) = 0 \right\} \tag{3.83}$$

where "sup" denotes the supremum of the argument and is formally defined as the smallest upper bound with respect to r for which the condition to the right of the semicolon is true. Let us prove the relation between (3.82) and (3.83). We start with the condition

$$\lim_{\sigma \to 0} P_e(\sigma, k)/Q\left(\frac{\sqrt{r}A_k}{\sigma}\right) = 0 \tag{3.84}$$

where $P_e(\sigma, k)$ is given by (3.76). From (A.4) in appendix A, we can determine the following

$$\begin{aligned} \left[\sqrt{r}A_k\right]^+ &< \sqrt{e_k(\sigma)} \\ \sqrt{r}A_k &< \sqrt{e_k(\sigma)} \\ r &< \frac{e_k(\sigma)}{A_k^2}, \end{aligned} \tag{3.85}$$

where the operation $[.]^+$ chooses either zero or the argument, depending on which is the larger of the two. From (3.78), and since e_k and A_k^2 can only be positive, r can take a value between zero and one. We can now make the right side of the inequality (3.85) a minimum upper bound of r by taking the limit $\sigma \to 0$:

$$r < \lim_{\sigma \to 0} \frac{e_k(\sigma)}{A_k^2}. \tag{3.86}$$

Since the right side of the inequality (3.86) is the minimum upper bound of r,

$$\sup \left\{ 0 \leq r \leq 1 : \lim_{\sigma \to 0} P_e(\sigma, k)/Q\left(\frac{\sqrt{r}A_k}{\sigma}\right) = 0 \right\} = \eta_k = \lim_{\sigma \to 0} \frac{e_k(\sigma)}{A_k^2}. \tag{3.87}$$

An equivalent expression for η_k in [31] is

$$\eta_k = \frac{2}{A_k^2} \lim_{\sigma \to 0} \sigma^2 \log 1/P_e(\sigma, k). \tag{3.88}$$

From (3.88) it can be concluded that in the situations where the BEP does not approach zero as $\sigma \to 0$, such as the single user closed eye situation, the multiuser efficiency is 0. On the other hand, if the

multiuser efficiency is a positive value, the bit error rate approaches zero exponentially as $\sigma \rightarrow 0$. The multiuser efficiency is very close to the asymptotic multiuser efficiency, unless the SNR is very low.

Verdu [31] defines the *worst asymptotic effective energy* ω as the minimum effective energy among all users as $\sigma \rightarrow 0$. That is

$$\omega(A_1, \ldots, A_K) \stackrel{\text{def}}{=} \min_{k=1,\ldots,K} \lim_{\sigma \rightarrow 0} e_k(\sigma) \tag{3.89}$$

$$= \min_{k=1,\ldots,K} \lim_{\sigma \rightarrow 0} A_k^2 \eta_k \tag{3.90}$$

$$= 2 \lim_{\sigma \rightarrow 0} \sigma^2 \log 1/P \left[\bigcup_{k=1}^{K} \{b_k \neq \hat{b}_k\} \right] \tag{3.91}$$

provided that $\omega(A_1, \ldots, A_K) > 0$. Equation (3.90) follows from (3.82), and (3.91) follows from taking the $\lim_{\sigma \rightarrow 0} \sigma^2 \log(\cdot)$ of both sides of

$$\max_{k=1,\ldots,K} P_e(\sigma, k) \leq P \left[\bigcup_{k=1}^{K} \{b_k \neq \hat{b}_k\} \right] \leq \sum_{k=1}^{K} P_e(\sigma, k). \tag{3.92}$$

The *near-far resistance* [46] is a figure of merit which defines the detector in terms of the near-far capture immunity and is defined as the minimum asymptotic efficiency over the received energies of all the other users, i.e.

$$\bar{\eta}_k = \inf_{\substack{A_j > 0 \\ j \neq k}} \eta_k, \tag{3.93}$$

where "inf" denotes the infimum, and is defined as the maximum lower bound of the argument. In the case where we have received energies which vary with time (such a the mobile channel), we have a more restrictive definition

$$\bar{\eta}_k = \inf_{\substack{A_j[i] > 0 \\ (i,j) \neq (0,k)}} \eta_k. \tag{3.94}$$

3.5.1 ASYMPTOTIC MULTIUSER EFFICIENCY OF THE TWO USER MATCHED FILTER

Let us consider asymptotic multiuser efficiency in terms of the two user matched filter case. For the matched filter receiver we have the case of the closed eye under the condition

$$A_1 \leq A_2 |\rho|, \tag{3.95}$$

where we have from equation (3.88) that if the BEP does not approach zero as $\sigma \to 0$ the asymptotic multiuser efficiency is

$$\eta_k = 0 \quad \text{if } A_1 \leq A_2 \, |\rho|. \tag{3.96}$$

Conversely, in the open eye condition if

$$A_1 > A_2|\rho|, \tag{3.97}$$

we have from (A.3) and (3.35)

$$\begin{aligned} \lim_{\sigma \to 0} \frac{P_e(\sigma,1)}{Q\left(\frac{\sqrt{r}A_1}{\sigma}\right)} &= \lim_{\sigma \to 0} \frac{\frac{1}{2}Q\left(\frac{A_1 - A_2\rho}{\sigma}\right) + \frac{1}{2}Q\left(\frac{A_1 + A_2\rho}{\sigma}\right)}{Q\left(\frac{\sqrt{r}A_1}{\sigma}\right)} \\ &= 0, \quad \sqrt{r}A_1 < A_1 - A_2|\rho|, \end{aligned} \tag{3.98}$$

from which we get

$$r < \left(1 - \frac{A_2}{A_1}|\rho|\right)^2. \tag{3.99}$$

Taking $\sup\{r\}$, we get

$$\eta_1 = \left(1 - \frac{A_2}{A_1}|\rho|\right)^2. \tag{3.100}$$

Combining the asymptotic multiuser efficiency for both regions of (3.95) and (3.97), we get

$$\eta_1 = \left[\max\left\{0, 1 - \frac{A_2}{A_1}|\rho|\right\}\right]^2, \tag{3.101}$$

which is the asymptotic efficiency for the two user matched filter receiver. A linear plot of the asymptotic multiuser efficiency for the two user matched filter detector is given in Figure 3.14.

3.5.2 ASYMPTOTIC MULTIUSER EFFICIENCY OF THE K USER MATCHED FILTER

It is trivial to expand the expression for matched filter asymptotic multiuser efficiency to the K user case. Using the same reasoning as before, we can combine (A.3) and (3.55) to obtain

$$\eta_k = \left[\max\left\{0, 1 - \sum_{j \neq k} \frac{A_j}{A_k}|\rho_{jk}|\right\}\right]^2 \tag{3.102}$$

for the synchronous case, and combine (A.3) and (3.71) to obtain

$$\eta_k = \left[\max\left\{0, 1 - \sum_{j \neq k} \frac{A_j}{A_k}\left(|\rho_{jk}| + |\rho_{kj}|\right)\right\}\right]^2 \tag{3.103}$$

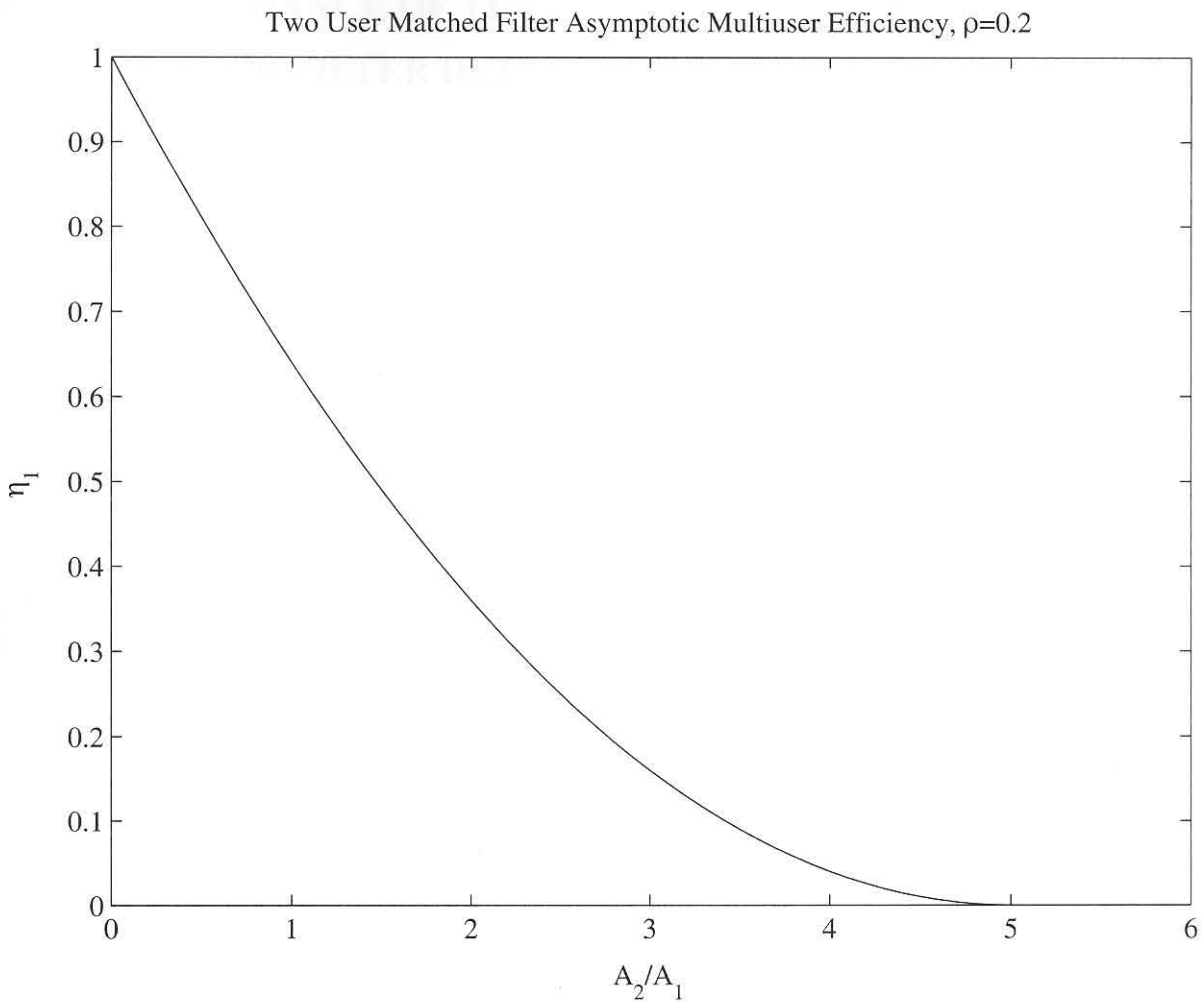

Figure 3.14: Asymptotic multiuser efficiency for a matched filter detector with two equal energy users and $\rho = 0.2$

for the asynchronous case. The asymptotic multiuser efficiency can be viewed as a normalized measure of the openness of the eye (refer to equations (3.57) and (3.72)).

Form (3.102) and (3.103) we can see that minimizing the asymptotic multiuser efficiency over all users, the near-far resistance of user k is equal to zero unless $\rho_{jk} = \rho_{kj} = 0 \; \forall \; j \neq k$ for all overlapping user bits for both the synchronous and asynchronous cases. This means that the single-user matched filter detector is not near-far resistant, since it is impossible that the orthogonality constraint can be maintained over all offsets in the asynchronous channel.

3.6 PERFORMANCE OF THE COHERENT SINGLE USER MATCHED FILTER DETECTOR IN FREQUENCY FLAT FADING

Let us now evaluate the single user matched filter where the received signals are subject to frequency flat Rayleigh fading. We assume coherent detection, i.e. that the fading amplitude and phase of the user of interest is perfectly known at the receiver. Let us adopt the synchronous version of the complex valued model in (2.31):

$$y(t) = \sum_{k=1}^{K} \sum_{i=1}^{M} \tilde{A}_k b_k[i] s_k(t - iT) + \sigma n(t), \tag{3.104}$$

where $\tilde{A}_k$ is the complex valued amplitude of user k due to a phase θ_k. The real and imaginary part of $\tilde{A}_k$ are independent and Gaussian with zero mean and standard deviation equal to A_k. First, let us examine the single user case.

3.6.1 THE SINGLE USER CASE IN THE PRESENCE OF FADING

Here we consider the one-shot model without loss of optimality, since we consider the fading coefficients to be perfectly known. In the single-user case we have

$$y(t) = \tilde{A}bs(t) + \sigma n(t), \quad t \in [0, T]. \tag{3.105}$$

The optimum decision rule selects the value of $b =\in \{\pm 1\}$ that minimizes the mean-square distance

$$\int_0^T \left| y(t) - \tilde{A}bs(t) \right|^2 dt = \int_0^T |y(t)|^2 \, dt + \int_0^T \left| \tilde{A}bs(t) \right| - 2\Re \left\{ \int_0^T y^*(t) \tilde{A}bs(t) dt \right\}, \tag{3.106}$$

that is, where the optimum decision rule is given by

$$\hat{b} = \operatorname{sgn} \left(\Re \left\{ \tilde{A} \int_0^T y^*(t) s(t) dt \right\} \right). \tag{3.107}$$

The inner product

$$y = \langle y^*, s \rangle = \int_0^T y^*(t) s(t) dt \tag{3.108}$$

is a sufficient statistic. The decision rule in (3.107) is equal to b only if the angle between the complex values $\tilde{A}$ and y is acute. That is to say that their absolute phase difference is less than $\pi/2$. Let us find the error probability of the decision rule by conditioning on the transmitted bits and the received fading coefficients:

$$\begin{aligned} P[\hat{b}=1|b=-1,\tilde{A}] &= P\left[-\left|\tilde{A}\right|^2+\sigma\Re\left\{\tilde{A}\int_0^T n^*(t)s(t)dt\right\}>0\,\middle|\,\tilde{A}\right] \\ &= P\left[-\left|\tilde{A}\right|^2+\sigma\Re\left\{\tilde{A}\right\}N_{\Re}+\sigma\Im\left\{\tilde{A}\right\}N_{\Im}>0\,\middle|\,\tilde{A}\right] \\ &= Q\left(\frac{\left|\tilde{A}\right|}{\sigma}\right), \end{aligned} \tag{3.109}$$

where

$$N_{\Re} \stackrel{\text{def}}{=} \Re\left\{\int_0^T n^*(t)s(t)dt\right\}, \tag{3.110}$$

$$N_{\Im} \stackrel{\text{def}}{=} \Im\left\{\int_0^T n^*(t)s(t)dt\right\}. \tag{3.111}$$

The symbols (3.110) and (3.111) denote Gaussian random variables with zero mean and unit variance. As in previous cases, the probability of error if a one is sent is identical. We assume that the received complex amplitude $\tilde{A}$ has a independent Rayleigh distributed real and imaginary parts. From (2.49) we have

$$f_R(r) = \begin{cases} r\exp\left(-\frac{r^2}{2}\right), & 0 \leq r \leq \infty, \\ 0, & r<0, \end{cases} \tag{3.112}$$

where R is the Rayleigh distributed random variable. We may write the received amplitude as the product of the Rayleigh distributed random variable and a deterministic part where $\tilde{A} = AR$. To find the BEP of the single user, we have to average over all values of the Rayleigh faded received amplitude. Subsequently, the BEP is given by

$$\begin{aligned} P_e^F(\sigma) &= E\left[Q\left(\frac{\left|\tilde{A}\right|}{\sigma}\right)\right] \\ &= \int_0^{\infty} r\exp\left(-\frac{r^2}{2}\right)Q\left(\frac{Ar}{\sigma}\right)dr \\ &= \frac{1}{2}\left(1-\frac{1}{\sqrt{1+\sigma^2/A^2}}\right), \end{aligned} \tag{3.113}$$

where (3.113) follows from (A.7). The BEP exhibits an interesting property when compared to the case of a deterministic amplitude. In the deterministic case, the decay in BEP is exponential. In the Rayleigh faded case, however, the BEP has a much slower hyperbolic decay. This highlights the detrimental effect a Rayleigh fading channel has on a digital communication system, and in our case,

a single user CDMA system.

The exact BEP of user k in the case of Rayleigh fading is given in [31] as

$$P_e^F(\sigma,k) = \frac{1}{2^{K-1}} \sum_{(b_1,\ldots,b_k)=(\{-1,1\},\ldots,\{-1,1\})} E\left[Q\left(\frac{\left|\tilde{A}_k\right|}{\sigma} + \sum_{j\neq k} b_j \frac{\Re\left\{\tilde{A}_j^* \tilde{A}_k\right\}}{\sigma\left|\tilde{A}_k\right|}\rho_{jk}\right)\right]$$

$$= E\left[Q\left(\frac{\left|\tilde{A}_k\right|}{\sigma} + \sum_{j\neq k} \frac{\Re\left\{\tilde{A}_j\right\}}{\sigma}\rho_{jk}\right)\right] \tag{3.114}$$

$$= E\left[Q\left(\frac{\left|\tilde{A}_k\right|}{\sqrt{\sigma^2 + \sum_{j\neq k} A_j^2\rho_{jk}}}\right)\right] \tag{3.115}$$

$$= \frac{1}{2}\left(1 - \frac{A_k}{\sqrt{\sigma^2 + \sum_{j} A_j^2\rho_{jk}}}\right), \tag{3.116}$$

where the phase term $\tilde{A}_k/\left|\tilde{A}_k\right|$ and the binary coefficients b_k have been dropped in (3.114) since they do not affect the distribution of the random variable inside the Q-function. Similar to the single user case, (3.115) follows from (A.7) because $\Re\{\tilde{A}_j\}$ are independent Gaussian random variables. We can obtain (3.116) by solving the averaging integral that led to the single user result. The asymptotic multiuser efficiency in the case of Rayleigh fading is given by [31],

$$\eta_k^F = \lim_{\sigma\to 0} \frac{\sigma^2}{4A_k^2 P_e^F(\sigma,k)}. \tag{3.117}$$

3.7 SUMMARY

The chapter begins by declaring the multiuser detection problem as a hypothesis testing problem. The concept of sufficient statistic is visited, and it is shown that the single user matched filter receiver contains sufficient statistic to make an optimal decision. The optimal (matched filter) single user receiver is analyzed and discussed. The CDMA matched filter detector for multiple users is presented, and is analyzed for the two user case. Performance measures such as BEP and power tradeoff regions are introduced, with the two user channel in mind. The phenomenon of the near-far effect is discussed as a basic limitation of the matched filter CDMA receiver. A useful visualization of the two user matched filter detector is presented in terms of a signal space representation.

The K user matched filter detection case is also analyzed in this chapter. The exact and Gaussian approximated BEP equations are derived and presented as a performance measure for the K user

case. The infinite user limit for BEP in a CDMA channel is also visited. Asymptotic multiuser efficiency and related measures such as near-far resistance and signal to interference ratios are also presented and discussed. The chapter is concluded with analysis of the matched filter detector in single and multiuser channels with frequency flat fading.

CHAPTER FOUR

LINEAR MULTIUSER DETECTORS

The matched filter detector is a linear detector. We will now examine other linear detectors (i.e. detectors that operate on the received samples by means of an arbitrary linear transformation $\mathbf{M}$ as shown in Figure 4.1). The class of linear multiuser detectors discussed in this chapter include the decorrelating detector, the MMSE detector, and the generalized extension of the aforementioned detectors: the optimum linear multiuser detector.

The blind detectors discussed in this dissertation all have mean weight vector solutions that converge to the MMSE solution. It is thus imperative to understand the operation of the MMSE detector, so we can make meaningful comparisons between the blind detectors and the MMSE detector, especially where multipath combining is concerned. In this chapter we will briefly visit the decorrelating detector, after which we will consider the optimum linear and MMSE detectors in more detail.

The performance of the MMSE detector is evaluated by means of performance measures presented in Chapter 3. An extension of the MMSE detector model provided in [31] to the multipath case is also presented. This is done by partially utilizing the derivation in [47].

4.1 THE LINEAR DECORRELATING DETECTOR

Before we discuss the optimum linear and MMSE detectors, let us briefly and qualitatively consider the operation of the linear decorrelating detector. The decorrelating detector is relevant to a certain extent, since the MMSE detector and decorrelating detector perform the same linear transformation when noise is absent from the channel [46], [31]. This means that both detectors exhibit the same asymptotic multiuser efficiency, and both are optimally near-far resistant. The CDMA decorrelating detector first proposed by Schneider [48] is equivalent to the zero-forcing equalizer, as it attempts to

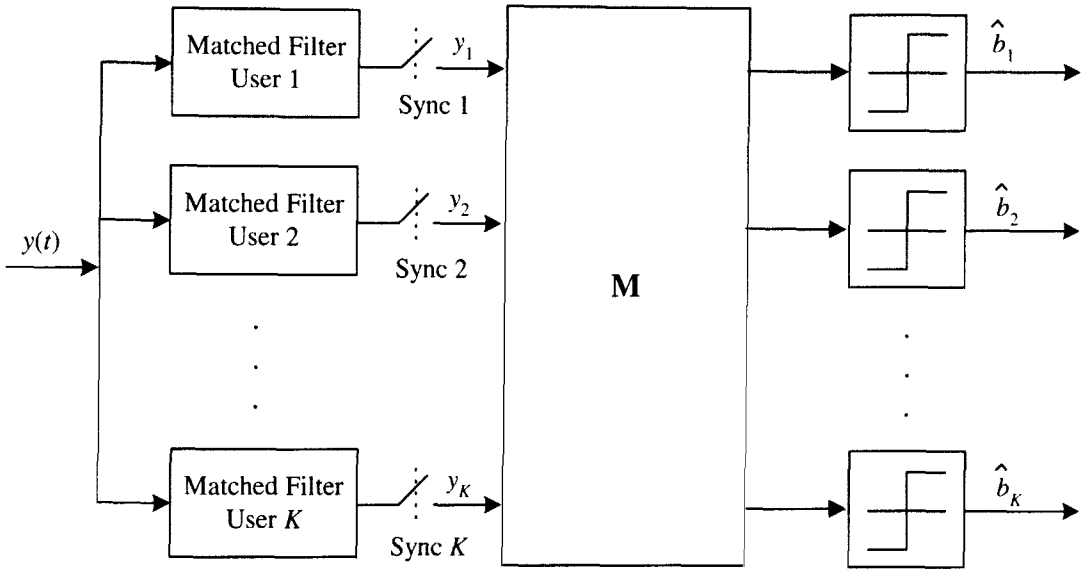

Figure 4.1: Block diagram depicting the structure of the K user linear receiver performing a linear operation $\mathbf{M}$ on the sampled matched filter outputs.

perform a linear inverse correlation matrix operation $\mathbf{M} = \mathbf{R}^{-1}$ on the received signal samples. In some cases, the correlation matrix may be singular, in which case a simple matrix inversion is not possible. A generalized inverse may then be used [46]. Analogous to zero-forcing equalization, noise enhancement may be a problem in the CDMA decorrelating detector. After Schneider, there have been several efforts to realize the decorrelating detector adaptively [49] [50].

Having now briefly visited the decorrelating detector, let us consider a generalized extension of the decorrelating detector.

4.2 THE OPTIMUM LINEAR DETECTOR

Lupas and Verdu [46] extended the MMSE and decorrelating detectors to the optimum linear detector. The class of linear detectors performs a linear transformation on the received signal vector. The optimum linear detector is the detector which maximizes the asymptotic multiuser efficiency for every vector of received amplitudes. In general, it is possible to achieve a certain tradeoff of interference rejection and attenuation of the desired signal component in order to maximize the asymptotic multiuser efficiency within the constraint of linear multiuser detection. Employing the complex vector matrix model of (2.17), let us denote the kth user linear transformation by $\mathbf{t_k}$, with

$$\hat{b}_k = \text{sgn}\left(\mathbf{t}_k^H \mathbf{y}\right), \tag{4.1}$$

where $\mathbf{y}$ is the complex vector of normalized matched filter outputs. Then

$$\mathbf{t}_k^H \mathbf{y} = \sum_{j=1}^{K} A_j b_j \mathbf{t}_k^H \mathbf{r}_j + \mathbf{t}_k^H \mathbf{n}, \tag{4.2}$$

where $\mathbf{r}_j$ is the jth column of the normalized crosscorrelation matrix $\mathbf{R}$. The probability of error achieved by the transformation $\mathbf{t}_k$ can be expressed as

$$P_e^{\mathbf{t}_k} = E\left[Q\left(\frac{A_k \mathbf{t}_k^H \mathbf{r}_k + \sum_{j \neq k} A_j b_j \mathbf{t}_k^H \mathbf{r}_j}{\sigma\sqrt{\mathbf{t}_k^H \mathbf{R} \mathbf{t}_k}}\right)\right], \tag{4.3}$$

where the expectation is with respect to b_j, $j \neq k$. The asymptotic multiuser efficiency of user k is given by the square of the smallest argument of the Q-function normalized by A_k^2/σ^2, i.e.

$$\eta_k(\mathbf{t}_k) = \frac{1}{\mathbf{t}_k^H \mathbf{R} \mathbf{t}_k}\left[\max\left\{0, \mathbf{t}_k^H \mathbf{r}_k - \sum_{j \neq k} \frac{A_j}{A_k} \left|\mathbf{t}_k^H \mathbf{r}_j\right|\right\}\right]^2. \tag{4.4}$$

Due to the presence of the absolute value in (4.4), the maximization of the K-user asymptotic multiuser efficiency entails solving a nonlinear optimization problem that does not permit a closed form solution. Lupas and Verdu [46] presented an algorithm to implement the kth user maximal linear asymptotic multiuser efficiency detector. The authors also presented sufficient conditions for the best linear detector to achieve optimum kth user multiuser efficiency, as well as sufficient conditions for decorrelating detector to be the best kth user linear detector. The computational complexity of the kth user maximal linear asymptotic multiuser efficiency detector is prohibitive for a large number of users when using the algorithm mentioned above.

Although it is not possible to find a closed form solution for the kth user asymptotic multiuser efficiency in a K user channel, it is possible, however, to evaluate a closed form solution in the two user case.

4.2.1 THE TWO USER OPTIMUM LINEAR DETECTOR

We will now examine the optimization of (4.4) with respect to $\mathbf{t}_k^H$ by analyzing the two-user case. As in Chapter 3, we will restrict ourselves to the real domain, as it is instrumental in understanding and visualizing the two user linear case. Without loss of generality, if we let $\mathbf{t}_1 = [1 \;\; x]^H$, the asymptotic multiuser efficiency becomes

$$\eta_1(\mathbf{t}_1) = \left[\max\left\{0, \frac{1 + x\rho - \frac{A_2}{A_1}|x + \rho|}{\sqrt{1 + 2\rho x + x^2}}\right\}\right]^2 = \left[\max\left\{0, g\left(x, \rho, \frac{A_2}{A_1}\right)\right\}\right]^2 \tag{4.5}$$

The value of x that maximizes $g\left(x, \rho, \frac{A_2}{A_1}\right)$ is

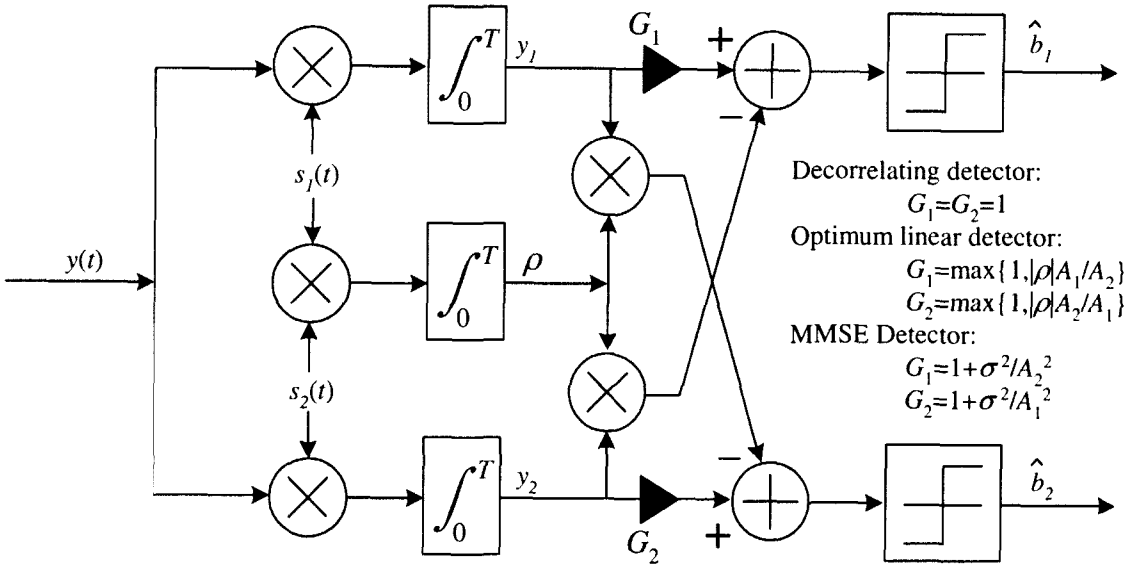

Figure 4.2: Block diagram depicting the structure of the two user linear decorrelating, optimum and MMSE receivers in the real domain.

$$\bar{x} = \begin{cases} -\frac{A_2}{A_1}\mathrm{sgn}(\rho) & \text{if } A_2/A_1 < |\rho| \\ -\rho & \text{otherwise} \end{cases} \tag{4.6}$$

When the relative energy of the interferer is strong enough, i.e. $A_2 \geq A_1|\rho|$, then the decorrelating detector maximizes asymptotic efficiency among all linear transformations. On the other hand, if $A_2 < A_1|\rho|$, then the received signal is correlated with

$$s_1(t) - \frac{A_2}{A_1}\mathrm{sgn}(\rho)s_2(t) \tag{4.7}$$

or equivalently with

$$\frac{A_1}{A_2}|\rho|s_1(t) - \rho s_2(t). \tag{4.8}$$

The optimum linear detector is a compromise solution between the decorrelating detector and the single user matched filter. As the relative power of the interferer decreases the optimum linear detector approaches the matched filter (Figure 4.2).

The maximum asymptotic multiuser efficiency for the two-user case is obtained by substituting (4.6) in (4.5).

$$\eta_1(\bar{\mathbf{t}}_1) = \begin{cases} 1 + \frac{A_2^2}{A_1^2} - 2|\rho|\frac{A_2}{A_1}, & \text{if } A_2/A_1 < |\rho| \\ 1 - \rho^2, & \text{otherwise} \end{cases} \tag{4.9}$$

When $A_2/A_1 < |\rho|$, the near-far resistance of the optimum linear detector is equal to that of the optimum multiuser detector in the high SNR region [46], [31]. On the other hand, when $A_2/A_1 \geq |\rho|$, there is no point, as far as near-far resistance is concerned, in utilizing the values of the received energies.

Note that the optimal linear asymptotic multiuser efficiency detector is optimal only in the high SNR region. It suffers from similar shortcomings as the decorrelating detector with respect to noise enhancement. In a low SNR environment, the single-user matched filter outperforms the optimal linear asymptotic multiuser efficiency detector. It is evident that there is room for improvement, if the noise is taken into account, concerning the performance of the optimal linear asymptotic multiuser efficiency detector. On the one hand, we have the matched filter detector that is optimized for white Gaussian noise. On the other hand, the decorrelating detector mitigates multiuser interference while disregarding the white Gaussian noise. The detector that utilizes information concerning both the SNR and MAI is the MMSE detector.

4.3 THE LINEAR MMSE DETECTOR

The adaptive MMSE detector [51], [52], [53] may solve many of the complexity and assumed knowledge problems associated with many of the other multiuser detector structures. As with matched filtering and de-correlation, the MMSE detection is a linear operation. This has the advantage that the received signal samples can be processed directly, thus simultaneously performing both the function of matched filtering and multiuser detection [7]. The MMSE detector turns the problem of multiuser detection into a problem of linear estimation. This is accomplished by minimizing a mean square error (MSE) cost criterion adaptively. The minimization can be done collectively over all users, or for each user individually.

An important quality of the MMSE detector, is that in addition to multiuser interference cancellation, it can also perform multipath (diversity) combining [54], [47], providing it has adequate filter span and that the channel inverse can be accurately modelled by a finite linear filter. In addition, the MMSE detector is successful at simultaneously mitigating narrow band interference (NBI) and MAI [55], [56]. A drawback of the MMSE detector is that a training sequence is needed to initially determine the CDMA channel conditions. After initial training, the MMSE detector can switch to its own decisions from which the MSE can be determined. This is referred to as *decision directed* mode. The imposition of training sequences implies some system overhead in the form of preamble and midamble bit sequences. The only knowledge required by the receiver is the training sequence of the user of interest. This means that the MMSE detector can be seen as a *single-user detector* capable of *multiuser interference cancellation* [7].

4.3.1 THE MMSE OPTIMIZATION PROBLEM

We start by quantitatively discussing the MMSE detector in terms of the MMSE optimization problem. Note that the same notation is used as in Chapters 2 and 3. The kth user MMSE detector chooses a complex waveform (or linear transformation) c_k of duration T that performs

$$\min_{c_k} E\left[\left(b_k - \langle y, c_k^* \rangle\right)\left(b_k - \langle y, c_k^* \rangle\right)^*\right] \tag{4.10}$$

and makes the decision

$$\hat{b}_k = \operatorname{sgn}\left(\langle y, c_k^* \rangle\right) \tag{4.11}$$

The MMSE linear transformation maximizes the SIR at the output of the linear transformation, i.e.

$$\frac{1}{\min\limits_{c_k} E\left[\left(b_k - \langle y, c_k^* \rangle\right)\left(b_k - \langle y, c_k^* \rangle\right)^*\right]} = 1 + \max_{c_k} \frac{E\left[\left(\langle A_k b_k s_k, c_k^* \rangle\right)\left(\langle A_k b_k s_k, c_k^* \rangle\right)^*\right]}{E\left[\left(\langle y - A_k b_k s_k, c_k^* \rangle\right)\left(\langle y - A_k b_k s_k, c_k^* \rangle\right)^*\right]}. \tag{4.12}$$

In orthogonal representation, we can always express c_k as

$$c_k = c_k^s + c_k^o, \tag{4.13}$$

where c_k^s is spanned by the signature waveforms $s_1, \ldots, s_K$ and c_k^o is orthogonal to the signature waveforms. Them we have

$$E\left[\left(b_k - \langle y, c_k^* \rangle\right)\left(b_k - \langle y, c_k^* \rangle\right)^*\right] = E\left[\left(b_k - \langle y, c_k^{s*} \rangle\right)\left(b_k - \langle y, c_k^{s*} \rangle\right)^*\right] + \sigma^2 \left\| c_k^o \right\|. \tag{4.14}$$

We will restrict ourselves to c_k spanned by the signature waveforms, i.e. a weighted combination of the matched filter outputs.

To analyze the operation and performance of the MMSE detector, we will start by formulating the vector matrix model of the MMSE detector.

4.3.2 THE MMSE DETECTOR VECTOR MATRIX MODEL

Let us start with the complex vector matrix model defined in (2.17)

$$\mathbf{y} = \mathbf{RAb} + \mathbf{n}, \tag{4.15}$$

where $\mathbf{R}$ is the correlation matrix in Hermitian form, $\mathbf{A}$ is a complex diagonal matrix of the user amplitudes, and $\mathbf{n}$ is a complex valued Gaussian vector with independent real and imaginary components

and with a covariance matrix equal to $2\sigma^2\mathbf{R}$. The complex bit vector is denoted by $\mathbf{b}$. The MMSE detector attempts to minimized the MSE or the difference between the actual transmitted bit vector $\mathbf{b}$ and a linear complex transformation $\mathbf{M}$ of the received signal vector $\mathbf{y}$ by adjusting the transformation $\mathbf{M}$. The transformation $\mathbf{M}$ is a $K \times K$ matrix and a K user joint optimization problem. Equivalent to the joint optimization problem, we can also have K uncoupled optimization problems (one for each user), in which case the real error cost function of user k is given by the expected value of the squared error, i.e.

$$J_k = E\left[\left(b_k - \mathbf{m}_k^H \mathbf{y}\right)^2\right], \tag{4.16}$$

where $\mathbf{m_k}$ is the kth column vector of $\mathbf{M}$. In the K user joint optimization problem, the real error cost function J is given by

$$J = E\left[(\mathbf{b} - \mathbf{My})^H (\mathbf{b} - \mathbf{My})\right] = E\left[\mathbf{e}^H \mathbf{e}\right] \tag{4.17}$$

where $\mathbf{e}$ denotes the complex error vector. Alternatively, the real error cost function is given by the trace of the covariance matrix $\mathbf{J}$ of the error vector, i. e.

$$\begin{aligned} J &= \mathrm{tr}\{\mathbf{J}\} && (4.18)\\ &= \mathrm{tr}\left\{E\left[(\mathbf{b} - \mathbf{My})(\mathbf{b} - \mathbf{My})^H\right]\right\} && (4.19)\\ &= \mathrm{tr}\{E\left[\mathbf{e}\mathbf{e}^H\right]\}. && (4.20) \end{aligned}$$

To find the complex matrix $\mathbf{M}$ that will minimize the cost function J, we will use the gradient method. This is done by partially differentiating the cost function J with respect to the complex elements of $\mathbf{M}$, equating it to zero, and solving for $\mathbf{M}$. The matrix $\mathbf{M}$ has complex elements, which can be written in the form

$$m_{vw} = x_{vw} + jy_{vw}. \tag{4.21}$$

The definition for the element of the vth row and the wth column of the *complex gradient operator* [57] matrix $\nabla_{\mathbf{M}}$ is given by

$$\nabla_{m_{vw}} = \frac{\partial}{\partial x_{vw}} + j\frac{\partial}{\partial y_{vw}}. \tag{4.22}$$

To solve the MSE cost function optimization problem, we will first apply the gradient operator to the real cost function J. The *complex gradient matrix* $\nabla_{\mathbf{M}}(J)$ is thus given by

$$\nabla_{\mathbf{M}}(J) = \begin{bmatrix} \frac{\partial J}{\partial x_{11}} + j\frac{\partial J}{\partial y_{11}} & \cdots & \frac{\partial J}{\partial x_{1K}} + j\frac{\partial J}{\partial y_{1K}} \\ \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial x_{K1}} + j\frac{\partial J}{\partial y_{K1}} & \cdots & \frac{\partial J}{\partial x_{KK}} + j\frac{\partial J}{\partial y_{KK}} \end{bmatrix}, \tag{4.23}$$

where equation (4.23) represents a natural extension of the customary definition of a gradient for a function of real elements to the more general case of a function of complex elements.[1]
By letting $\nabla_{\mathbf{M}}(J) = \mathbf{0}$ and solving for $\mathbf{M}$, we will have found an expression for $\mathbf{M}$ where the error surface for each user in the K dimensional space has a minimum. To do this, let us first manipulate (4.14)

$$\begin{aligned} J &= \mathrm{tr}\left\{E\left[(\mathbf{b}-\mathbf{My})(\mathbf{b}-\mathbf{My})^H\right]\right\} \\ &= \mathrm{tr}\left\{E\left[(\mathbf{b}-\mathbf{My})(\mathbf{b}^H-\mathbf{y}^H\mathbf{M}^H)\right]\right\} \\ &= \mathrm{tr}\left\{E\left[\mathbf{bb}^H-\mathbf{by}^H\mathbf{M}^H-\mathbf{Myb}^H+\mathbf{Myy}^H\mathbf{M}^H\right]\right\} \\ &= \mathrm{tr}\left\{E\left[\mathbf{bb}^H\right]-E\left[\mathbf{by}^H\right]\mathbf{M}^H-\mathbf{M}E\left[\mathbf{yb}^H\right]+\mathbf{M}E\left[\mathbf{yy}^H\right]\mathbf{M}^H\right\}, \end{aligned} \tag{4.24}$$

where (4.24) follows from the fact that $\mathbf{M}$ is assumed to be constant. Assuming no correlation between the data of different users or between the data and noise vectors, we have

$$E\left[\mathbf{bb}^H\right] = 2\mathbf{I}, \tag{4.25}$$
$$E\left[\mathbf{by}^H\right] = 2\mathbf{AR}, \tag{4.26}$$
$$E\left[\mathbf{yb}^H\right] = 2\mathbf{RA}, \tag{4.27}$$
$$E\left[\mathbf{yy}^H\right] = 2\mathbf{RA}^2\mathbf{R}+2\sigma^2\mathbf{R}. \tag{4.28}$$

Simplifying the cost function with the above results, we have

$$\begin{aligned} J &= \mathrm{tr}\left\{2\mathbf{I}-2\mathbf{ARM}^H-2\mathbf{MAR}+2\mathbf{M}(\mathbf{RA}^2\mathbf{R}+\sigma^2\mathbf{R})\mathbf{M}^H\right\} \\ &= \mathrm{tr}\left\{2\mathbf{I}-2\mathbf{ARM}^H-2\mathbf{MAR}+2\mathbf{MRA}^2\mathbf{RM}^H+\mathbf{M}\sigma^2\mathbf{RM}^H\right\}. \end{aligned} \tag{4.29}$$

Let us now find the complex gradient matrix $\nabla_{\mathbf{M}}(J)$ of the cost function,

$$\begin{aligned} &\nabla_{\mathbf{M}}(J) \\ &\quad = 2\tfrac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\{2\mathbf{I}\}-\mathrm{tr}\left\{2\mathbf{ARM}^H\right\}-\mathrm{tr}\{2\mathbf{MAR}\}+\mathrm{tr}\left\{2\mathbf{MRA}^2\mathbf{RM}^H\right\}+\mathrm{tr}\left\{2\mathbf{M}\sigma^2\mathbf{RM}^H\right\}\right) \end{aligned} \tag{4.30}$$

where (4.30) is evaluated in Appendix C. The result of the gradient of the cost function J from Appendix C is given by

$$\nabla_{\mathbf{M}}(J) = -4\mathbf{AR}+4\mathbf{MRA}^2\mathbf{R}+4\mathbf{M}\sigma^2\mathbf{R}. \tag{4.31}$$

[1]Note that the cost function J is *not* analytic, when it is written in terms of complex filter taps. The definition of the derivative of the cost function J with respect to the complex transformation matrix $\mathbf{M}$ requires special attention. This issue is discussed in Appendix C where the relation between derivative and gradient with respect to a complex valued matrix is discussed.

To obtain the minimum on the error surface, the slope or gradient must be set equal to zero, i.e

$$\nabla_{\bar{\mathbf{M}}}(J) = -4\mathbf{AR} + 4\bar{\mathbf{M}}\mathbf{R}\mathbf{A}^2\mathbf{R} + 4\bar{\mathbf{M}}\sigma^2\mathbf{R} = \mathbf{0} \tag{4.32}$$

with $\bar{\mathbf{M}}$ the optimum value for the linear transformation $\mathbf{M}$. Solving for $\bar{\mathbf{M}}$ we obtain

$$\bar{\mathbf{M}} = \mathbf{A}^{-1}\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}. \tag{4.33}$$

The MMSE detector outputs the following decision for user k

$$\hat{b}_k = \operatorname{sgn}\left(\frac{1}{A_k}\left[\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}\mathbf{y}\right]_k\right) \tag{4.34}$$

$$= \operatorname{sgn}\left(\left[\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}\mathbf{y}\right]_k\right). \tag{4.35}$$

Note that the dependence of the MMSE detector on received amplitudes is only through the signal-to-noise ratios A_k^2/σ^2 due to the sgn function. Because of this, we can replace the optimum linear transformation in (4.34) with

$$\tilde{\mathbf{M}} = \left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}. \tag{4.36}$$

In the formulation of the MMSE detector vector matrix model, we have assumed a great deal less than in the basic CDMA model. We did not assume that the background noise is Gaussian, nor that the bits are binary valued. The only assumptions we made were that the bits were uncorrelated from user to user, that the bit and noise vectors were uncorrelated, and that $E[b_k^2] = 1$.

4.3.3 THE TWO USER MMSE DETECTOR

Once again restricting ourselves to the real domain, in the two user case we have from (4.36)

$$\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1} = \left[\left(1 + \frac{\sigma^2}{A_1^2}\right)\left(1 + \frac{\sigma^2}{A_2^2}\right) - \rho^2\right]^{-1}\begin{bmatrix} 1 + \frac{\sigma^2}{A_2^2} & -\rho \\ -\rho & 1 + \frac{\sigma^2}{A_1^2} \end{bmatrix} \tag{4.37}$$

from which the two user MMSE detector follows, as shown in Figure 4.2.

4.3.4 THE LIMITING FORMS OF THE MMSE DETECTOR

The MMSE detector is a compromise between the matched filter detector and the decorrelating detector. To illustrate this, we shall investigate the linear transformation $\mathbf{M}^* = (\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1}$ in its limiting forms as $\sigma \to 0$ and $\sigma \to \infty$. On the one hand, if $\sigma \to 0$, then $(\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1} \to \mathbf{R}^{-1}$, which means that the MMSE detector approaches the decorrelating detector. On the other hand, if $\sigma \to \infty$, the matrix $(\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1}$ becomes strongly diagonal, and the MMSE detector approaches

the conventional matched filter detector.

The above results reinforces the statement that the asymptotic multiuser efficiency and near-far resistance of the MMSE detector is equal to that of the decorrelating detector. This is intuitive, as the asymptotic multiuser efficiency and near-far resistance performance measures are evaluated in the limit as $\sigma \to 0$.

4.3.5 THE ASYNCHRONOUS MMSE DETECTOR

The linear time invariant transfer function of the asynchronous MMSE detector for a K user CDMA channel is given by

$$\tilde{\mathbf{M}}_a = \left(\mathbf{R}^H[1]z + \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} + \mathbf{R}[1]z^{-1}\right)^{-1}. \tag{4.38}$$

This is verified in [31] parallel to the asynchronous decorrelating detector, and is the limiting form of the inverse of the equivalent correlation matrix that we would obtain for a finite frame length (refer to (2.40)). The equivalent correlation matrix for a finite frame length is in the form

$$\mathbf{R}_{a,\text{MMSE}} = \begin{bmatrix} \mathbf{R}[0]+\sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] & 0 & \cdots & 0 \\ \mathbf{R}[1] & \mathbf{R}[0]+\sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] & \cdots & \vdots \\ 0 & \mathbf{R}[1] & \ddots & \vdots & 0 \\ \vdots & \vdots & \cdots & \mathbf{R}[0]+\sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] \\ 0 & \cdots & 0 & \mathbf{R}[1] & \mathbf{R}[0]+\sigma^2\mathbf{A}^{-2} \end{bmatrix}. \tag{4.39}$$

4.3.6 THE WIENER FILTER CHARACTERIZATION OF THE MMSE DETECTOR

For the Wiener filter characterization of the MMSE CDMA detector we return to the synchronous case. To illustrate the operation of the Wiener filter, we will use the model of orthonormal projections as in (2.23). We will limit ourselves to the uncoupled optimization problem, where optimization is done with respect to a single user. Without loss of generality, we consider user 1 as the desired user. We will start by defining a vector $\mathbf{p}$, which is the *cross correlation vector* between the vector $\mathbf{r}$ (2.25) and desired response b_1:

$$\mathbf{p} = E\left[b_1^*\mathbf{r}\right]. \tag{4.40}$$

The optimal vector transformation $\bar{\mathbf{v}}$ that minimizes the mean square error for user 1

$$E\left[e_1 e_1^*\right] = E\left[\left(b_1 - \mathbf{v}^H\mathbf{r}\right)\left(b_1^* - \mathbf{r}^H\mathbf{v}\right)\right] \tag{4.41}$$

can be obtained by setting the gradient equal to the zero-vector, i.e.

$$E\left[b_1^*\mathbf{r} - \mathbf{r}\mathbf{r}^H\mathbf{v}\right] = \mathbf{0}, \tag{4.42}$$

where the gradient of a complex vector is again defined as in the vector case of (C.9). The first term $\mathbf{p} = E[b_1^*\mathbf{r}]$ can be simplified to

$$\mathbf{p} = E\left[b_1^*\mathbf{r}\right] = 2A_1\mathbf{s}_1, \tag{4.43}$$

from (2.25) and from the fact that the noise and data is uncorrelated and also from the fact that the data of different users is uncorrelated. The second term $E[\mathbf{r}\mathbf{r}^H]$ was derived in (2.27) and is equal to

$$E[\mathbf{r}\mathbf{r}^H] = 2\sigma^2\mathbf{I} + 2\sum_{k=1}^{K} A_k^2\mathbf{s}_k\mathbf{s}_k^H. \tag{4.44}$$

Solving for $\mathbf{v}$, we obtain the optimum solution for the linear vector transform

$$\begin{aligned}
\bar{\mathbf{v}} &= \left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right] && (4.45)\\
&= \mathbf{C}^{-1}\mathbf{p} && (4.46)\\
&= A_1\left[\sigma^2\mathbf{I} + \sum_{k=1}^{K} A_k^2\mathbf{s}_k\mathbf{s}_k^H\right]^{-1}\mathbf{s}_1, && (4.47)
\end{aligned}$$

where $\mathbf{C}$ denotes the covariance matrix of the vector $\mathbf{r}$ divided by 2, and is given by

$$\mathbf{C} = \sigma^2\mathbf{I} + \sum_{k=1}^{K} A_k^2\mathbf{s}_k\mathbf{s}_k^H \tag{4.48}$$

Equation (4.46) is an expression of the *Wiener-Hopf* equation [57], [58]. It is beneficial to know the minimum mean-square error achievable with the detector depending on the channel noise. The MMSE is given by

$$
\begin{aligned}
J_{\min} &= E\left[\left(b_1 - \bar{\mathbf{v}}^H\mathbf{r}\right)\left(b_1^* - \mathbf{r}^H\bar{\mathbf{v}}\right)\right] \\
&= E\left[b_1 b_1^*\right] - E\left[b_1^*\bar{\mathbf{v}}^H\mathbf{r}\right] - E\left[b_1\mathbf{r}^H\bar{\mathbf{v}}\right] + E\left[\bar{\mathbf{v}}^H\mathbf{r}\mathbf{r}^H\bar{\mathbf{v}}\right] \\
&= 2 - \left\{\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right]\right\}^H E\left[\mathbf{r}\mathbf{r}^H\right]\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right] \qquad (4.49) \\
&= 2 - \left\{\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right]\right\}^H E\left[b_1^*\mathbf{r}\right] \\
&= 2 - E\left[b_1^*\mathbf{r}\right]^H\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right] \qquad (4.50) \\
&= 2 - \mathbf{p}^H\mathbf{C}^{-1}\mathbf{p} \qquad (4.51) \\
&= 2 - 2A_1^2\mathbf{s}_1^H\left[\sigma^2\mathbf{I} + \sum_{k=1}^{K} A_k^2\mathbf{s}_k\mathbf{s}_k^H\right]^{-1}\mathbf{s}_1 \qquad (4.52)
\end{aligned}
$$

The expression in (4.51) corresponds to the expression of minimum mean-squared error of the standard Wiener filter as evaluated in [57].

4.4 THE MMSE DETECTOR LEAST MEAN SQUARE (LMS) ALGORITHM

From equation (4.47) it can be seen that to determine the optimum solution, a matrix inversion needs to be performed. This is a computationally expensive operation, and other methods need to be considered to avoid this. In addition, mobile channels are time varying, and the detector needs to follow these variations. The LMS algorithm achieves the aforementioned by being simple to implement, being able to learn the channel impulse response adaptively, and being able to follow time channel variations. For correct operation of the LMS algorithm, high certainty data of the desired user must be available at the receiver. This seems like to much to ask, as the data is what we need to determine in the first place. However, this requirement can be fulfilled by sending a *training sequence* to insure initial convergence. After this, the demodulated bits have a high certainty, and can be used by the MMSE detector to follow variations in the channel. The latter mode is referred to as *decision directed* operation. In this way the MMSE detector can be adaptively implemented, but with the disadvantage of some overhead in the form of training sequences.
The operation of the LMS algorithm can be seen as a *feedback control system*. It consists of two basic processes [57], i.e.

- An *adaptive* process which involves the adaptation of the tap weights.
- A *filtering* process which involves the inner product of an input vector with the weight vector, as well as generating an estimation error which actuates the adaptive process.

The LMS algorithm is based on the method of *steepest descent*, which is one of the oldest methods of optimization. To find the minimum value of the mean squared error using the the steepest descent algorithm, we proceed as follows:

1. We begin with an initial value $\mathbf{v}[0]$ for the tap weight vector, which is an arbitrary value.
2. Using the initial or present guess, we compute the gradient vector, the real and the imaginary parts which are defined as the derivative of the mean-squared error $J[n]$, evaluated with respect to the real and imaginary parts of the tap weight vector $\mathbf{v}[n]$ at time n (or the nth iteration).
3. The next guess of the tap weight vector is computed by making a change in the initial or present guess in a direction opposite to that of the gradient vector.
4. Go back to step 2 and repeat.

If the cost function is convex, then the minimum will be found after several iterations of the above algorithm. The distance with which the next guess differs from the current guess is termed the *step size*.

Let us now examine the elements of stochastic gradient descent optimization. Suppose we wish to find the multi-dimensional parameter θ^* that minimizes the function

$$\Psi(\theta) = E\left[g\left(X, \theta\right)\right]. \tag{4.53}$$

For a step size μ, a convex function Ψ and a initial condition θ_0, it would be possible to converge to the global minimum via steepest descent

$$\theta_{j+1} = \theta_j - \mu \nabla \Psi(\theta_j). \tag{4.54}$$

If the step size is arbitrarily small, then eventually θ_j will be close enough to θ^* for all practical purposes. To speed up convergence, the step size can initially be large and progressively decreased as the algorithm converges. Other than the fact that Ψ is convex we did not invoked any structure in (4.53). In order to calculate the expected value, we need to know the distribution of X. This is not so in all cases in practice. Instead, let us assume that the algorithm is allowed to observed an independent sequence $\{X_1, X_2, \ldots\}$ where each of the random variables in the sequence has the same distribution as X. With this information we can estimate the distribution of X and also calculate an approximation to $\nabla\Psi$. This requires too much effort and a simpler approach would be to replace the expected value of the gradient by the immediate (noisy) gradient, i.e.

$$\Psi(\theta) = E\left[g\left(X, \theta\right)\right]. \tag{4.55}$$

This can be justified by the fact that although the immediate negative gradient does not necessarily point in the direction of steepest descent, the average negative gradient of a few iterations does. According to the *law of large numbers*, if the step size is infinitesimally small, the trajectory of the algorithm will very closely track the path of steepest descent. This algorithm is known as the stochastic gradient descent algorithm. In the case where the cost function is a quadratic error cost function, the stochastic gradient algorithm is known as the LMS algorithm. It is important to know that the stochastic gradient algorithm can also be used when the sequence of realizations of X is dependent, subject to the fact that the sequence is also ergodic (the time average of the immediate gradients converges to its expected value).

Applying the stochastic gradient algorithm to the MMSE case (LMS), the linear MMSE detector for user one correlates the received waveform with the signal c_1 that minimizes

$$E\left[(b_1 - \langle y, c_1^* \rangle)^2\right]. \tag{4.56}$$

How does this fit into the stochastic approximation framework that we have derived above? The function $g(X, c_1)$ is our mean square error cost, i.e.

$$g(X, c_1) = (b_1 - \langle y, c_1^* \rangle)^2, \tag{4.57}$$

where X represents the received waveform y and the bit b_1. It is easily verified that (4.57) is strictly convex in c_1. We first will consider the synchronous case, after which we will briefly address the asynchronous case. The independent identically distributed observations used in the stochastic gradient algorithm are $X_j = (b_1[j], y[j])$, where $y[j]$ is the received signal modulated by the jth bit of all the synchronous users. To specify the gradient algorithm of (4.54), all we need to do is evaluate the gradient of $(b_1 - \langle y, c_1^* \rangle)^2$ with respect to c_1, which is equal to

$$2\left(\langle y, c_1^* \rangle - b_1\right) y. \tag{4.58}$$

We thus conclude that, in practice, the update algorithm is simply

$$c_1[j] = c_1[j-1] - \mu\left(\langle y[j], c_1^*[j-1] \rangle - b_1[j]\right) y[j]. \tag{4.59}$$

Since in practice we are working with a finite dimensional vector implementation of the adaptive law, a few things need to be pointed out. If the signature waveforms are known, then the dimensionality of the adaptive vector need not be larger than K. We know that the MMSE receiver does not need to know the transmitted signature vectors. Fortunately, by using a finite dimensional basis known to span all received signature waveforms (such as chip-matched filters), there will be sufficient dimensionality to implement our linear adaptive LMS algorithm. It is furthermore sufficient to sample at the Nyquist

rate for approximately band limited chip waveforms in both the synchronous and asynchronous cases. Our LMS adaptation algorithm in finite vector form is then given by

$$\mathbf{v}_1[n] = \mathbf{v}_1[n-1] - \mu\left(\mathbf{v}^H[n-1]\mathbf{r}[n] - b_1[n]\right)\mathbf{r}[n]. \tag{4.60}$$

Global convergence of the LMS algorithm is shown in [31], subject to a sufficient decrease in step size as the algorithm progresses. The maximum step size to ensure convergence at any moment is given by

$$\mu_{\max} = \frac{2}{\sigma^2 + \lambda_{\max}}, \tag{4.61}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H$. To retain acceptable performance in the asynchronous case, we need to lengthen the observation window that spans more than one bit period. This implies that the inner product in the penalty function (4.56) is taken over the whole truncated window. This does not affect the convexity of the cost function, allowing the detector to converge to the MMSE solution.

It is expected that the detector will converge to the MMSE solution if the interference is constant. When these parameters are slowly time varying, it is still possible for an adaptive detector to follow these variations. In the case of a new user suddenly being powered on, the decisions might be unreliable in decision directed mode, and the desired user might not converge. In this case, the desired user will then request for the training sequence to be retransmitted. This implies more overhead, and is undesirable. It is for this reason that *blind* multiuser detectors (such as the constant modulus detector) warrant some investigating. Instead of using data (or decision directed) to adapt, the blind detectors utilize the cyclostationarity in the signature waveforms to minimize some given criterion.

4.5 PERFORMANCE OF THE MMSE DETECTOR

In this section we will consider the performance of the MMSE detector, using some of the measures in Chapter 3 to evaluate the detector.

4.5.1 SIGNAL-TO-INTERFERENCE RATIO OF THE MMSE DETECTOR

To derive the SIR of the MMSE detector, we start by defining the covariance matrix of the interference as

$$\mathbf{\Omega} \stackrel{\text{def}}{=} \sigma^2\mathbf{I} + \sum_{k=2}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H. \tag{4.62}$$

Note that user 1 is excluded from the sum. We can now write the optimum MMSE transformation of (4.47) and the MMSE of (4.52) as

$$\bar{\mathbf{v}} = \frac{A_1}{1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1} \boldsymbol{\Omega}^{-1} \mathbf{s}_1 \tag{4.63}$$

and

$$J_{\min} = \frac{2}{1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1}. \tag{4.64}$$

These two results follow from the fact that

$$\left[\boldsymbol{\Omega} + A_1^2 \mathbf{s}_1 \mathbf{s}_1^H\right]^{-1} = \left[1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1\right]^{-1} \boldsymbol{\Omega}^{-1}, \tag{4.65}$$

which can be proven using the *matrix inversion lemma* [57] or also known as *Woodbury's identity* [7].[2] Using the above results, and remembering that $\bar{\mathbf{v}}$ achieves the maximum output signal to interference ratio of all linear detectors, the SIR of user 1 can be written as

$$\begin{aligned}
\gamma_{c1} &= \frac{E\left[\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)^H\right]}{E\left[\left(\bar{\mathbf{v}}^H\left(\mathbf{r} - A_1 b_1 \mathbf{s}_1\right)\right)\left(\bar{\mathbf{v}}^H\left(\mathbf{r} - A_1 b_1 \mathbf{s}_1\right)\right)^H\right]} && (4.66)\\
&= \frac{E\left[\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(A_1^* b_1^* \mathbf{s}_1^H \bar{\mathbf{v}}\right)\right]}{E\left[\left(\bar{\mathbf{v}}^H \mathbf{r} - A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(\mathbf{r}^H \bar{\mathbf{v}} - A_1^* b_1^* \mathbf{s}_1^H \bar{\mathbf{v}}\right)\right]} \\
&= \frac{2}{E\left[\left(b_1 - \bar{\mathbf{v}}^H \mathbf{r}\right)\left(b_1^* - \mathbf{r}^H \bar{\mathbf{v}}\right)\right]} - 1 && (4.67)\\
&= A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1, && (4.68)
\end{aligned}$$

From (4.66), it can be seen that the SIR is the expectation of the squared linear transformation of the desired user contribution divided by the expectation of the squared linear transformation of the interferers' contribution.

4.5.2 ASYMPTOTIC MULTIUSER EFFICIENCY AND NEAR-FAR RESISTANCE OF THE MMSE DETECTOR

Since the operation of the decorrelating detector and that of the MMSE detector are identical in a noiseless environment, they have the same asymptotic multiuser efficiency and near-far resistance. The asymptotic multiuser efficiency of the MMSE (or decorrelating) detector is given in [46] by

[2] The matrix inversion lemma states that for positive definite square matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ related by $\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^H$, the inverse of $\mathbf{A}$ is given by $\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^H\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^H\mathbf{B}$.

$$\bar{\eta}_k = 1 - \mathbf{a}_k^H \mathbf{R}_k^+ \mathbf{a}_k \tag{4.69}$$

$$= \frac{1}{R_{kk}^+}, \tag{4.70}$$

where $\mathbf{R}^+$ is the *Moore-Penrose* generalized inverse and denotes the inverse of a singular (or non-singular) square matrix $\mathbf{R}$.[3] The subscript k of $\mathbf{R}_k$ denotes the removal of the kth row and the kth column from the matrix $\mathbf{R}$. The vector $\mathbf{a}_k$ is the kth column of $\mathbf{R}$ with the kth entry removed and contains the correlations between the kth user and all other users. The value R_{kk}^+ is the element of the kth row and kth column of the generalized inverse of $\mathbf{R}$.

From Figure (4.3), it is evident that the asymptotic multiuser efficiency defined in (4.70) does not depend on the amplitude of the interfering user. This implies that the MMSE detector's asymptotic multiuser efficiency and near-far resistance are all exactly equal.
In the asynchronous case, Verdu [31] shows the near-far resistance to be

$$\bar{\eta}_k = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\mathbf{R}^H[1] e^{j\omega} + \mathbf{R}[0] + \mathbf{R}[1] e^{-j\omega} \right]_{kk}^{+} d\omega \right)^{-1}. \tag{4.71}$$

Lupas and Verdu showed in [46] that the near-far resistance of the MMSE, optimum linear and de-correlating detector is equal to that of the optimum (non-linear) multiuser detector if the desired user is linearly independent from the other users.

4.5.3 BEP OF THE MMSE DETECTOR

The decorrelating detector is only an optimization with respect to interference, whereas the MMSE detector is an optimization with respect to the combined contribution of noise and interference. This effectively means that the MMSE transformation will inevitably allow some residual multiuser interference to remain. The consequence of this is that the derivation of the MMSE detector BEP is similar to that of the single user matched filter. As in the case of the single user matched filter, the decision statistic depends on the sum of a Gaussian random variable (due to AWGN) and a binomial random variable (due to residual multiple access interference). In the synchronous case, the first user MMSE decision statistic can be written as

$$\left(\tilde{\mathbf{M}}\mathbf{y}\right)_1 = \left(\left(\mathbf{R} + \sigma^2 \mathbf{A}^{-2} \right)^{-1} \mathbf{y} \right) \tag{4.72}$$

$$= B_1 \left(b_1 + \sum_{k=2}^{K} \beta_k b_k \right) + \sigma \hat{n}_1, \tag{4.73}$$

[3] A generalized inverse $\mathbf{C}$ of a matrix $\mathbf{B}$ is any matrix that satisfies: $\mathbf{CBC} = \mathbf{C}$ and $\mathbf{BCB} = \mathbf{B}$. The Moore-Penrose generalized inverse is the unique inverse for which $\mathbf{BC}$ and $\mathbf{CB}$ are symmetric. It follows that if $\mathbf{B}$ is a square non-singular matrix, then its Moore-Penrose generalized inverse is $\mathbf{B}^{-1}$.

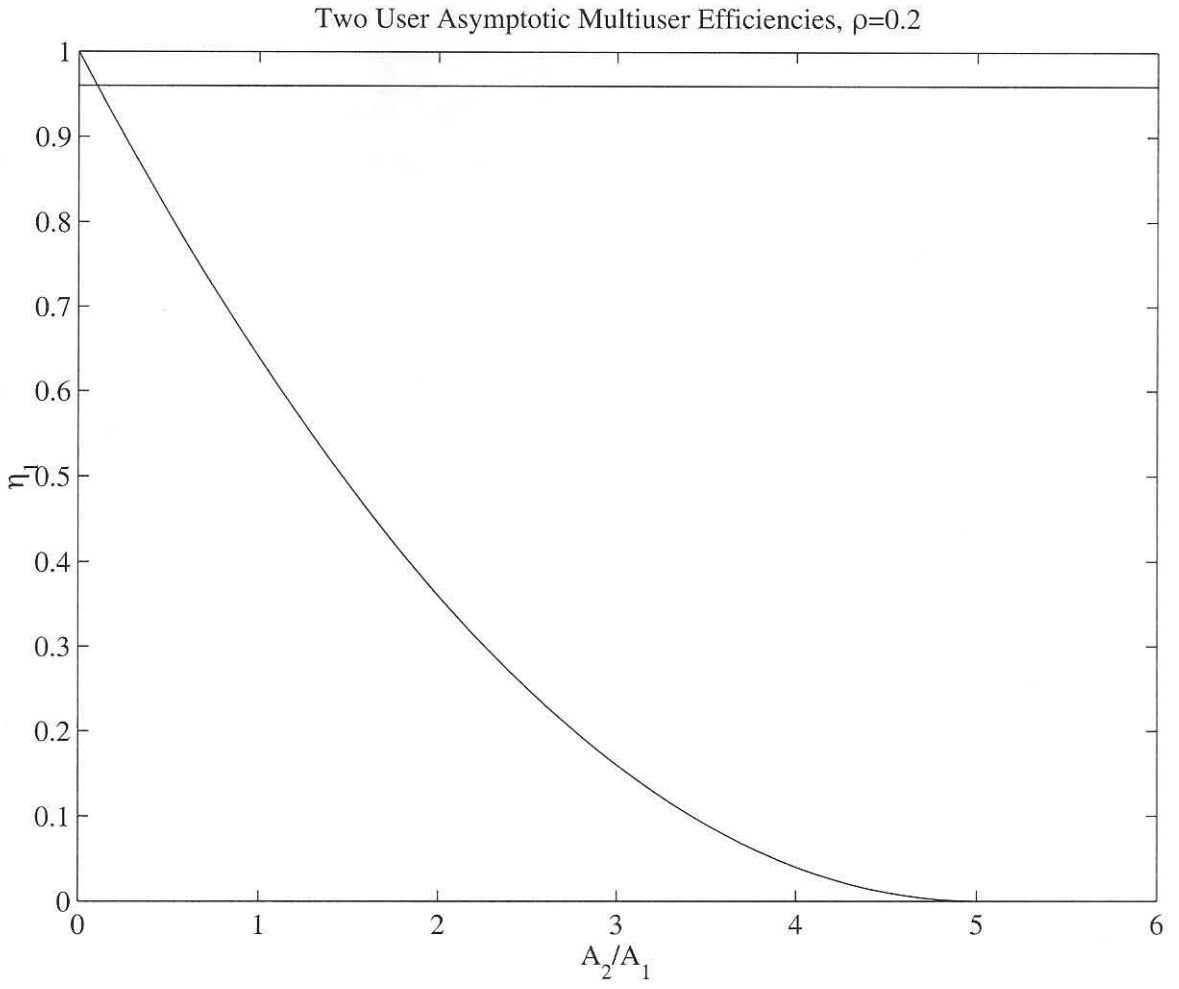

Figure 4.3: Asymptotic Multiuser Efficiencies of the Matched Filter, Decorrelating and MMSE Detectors.

where

$$B_k = A_k \left(\tilde{\mathbf{M}}\mathbf{R}\right)_{1k}, \tag{4.74}$$

$$\beta_k = \frac{B_k}{B_1}, \tag{4.75}$$

$$\hat{n}_1 \sim \mathcal{N}\left(0, \left(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}}\right)_{11}\right). \tag{4.76}$$

The symbol β_k denotes a measure of the residual interference of the kth interferer, and is termed the *leakage coefficient*. The Gaussian noise random variable is denoted by $\hat{n}_1$ and the binomial random variable is denoted by the sum in (4.73). The probability of error is given by

$$P_e(\sigma,1) = 2^{1-K} \sum_{b_1,\ldots,b_K \in \{-1,1\}^{K-1}} Q\left(\frac{A_1}{\sigma}\frac{(\tilde{\mathbf{M}}\mathbf{R})_{11}}{\sqrt{(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}})_{\mathbf{11}}}}\left(1+\sum_{k=2}^{K}\beta_k b_k\right)\right). \tag{4.77}$$

We face a similar problem as in the case of exact computation of the single user BEP, in that the number of computations grow exponentially with the number of active users. This is further complicated by the computation of the leakage coefficients. We will now apply the Gaussian approximation method to the MMSE case.

4.5.3.1 GAUSSIAN APPROXIMATION OF THE MMSE DETECTOR BEP

The Gaussian approximation method is surprisingly accurate when applied to the BEP of the MMSE detector. This is done by replacing the multiple access interference by a Gaussian random variable with identical variance, i.e. $Q(\mathrm{SIR}_1)$. We can use (4.68) together with (A.6) in Appendix A:

$$E\left[Q\left(\mu+\lambda X\right)\right] = Q\left(\frac{\mu}{\sqrt{1+\lambda^2}}\right), \tag{4.78}$$

where X is unit normal,

$$\mu = \frac{A_1}{\sigma}\frac{(\tilde{\mathbf{M}}\mathbf{R})_{11}}{\sqrt{(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}})_{\mathbf{11}}}} \tag{4.79}$$

and

$$\lambda^2 = \mu^2 \sum_{k=2}^{K} \beta_k. \tag{4.80}$$

Let us verify the accuracy of the approximation on an intuitive basis. We will qualitatively evaluate the deviation from Gaussianity of the decision statistic for the two limiting cases of $\sigma \to 0$ and $\sigma \to \infty$. As $\sigma \to 0$, the leakage coefficients disappear, removing the contribution of the binomial random variable. On the other hand, as $\sigma \to \infty$, the Gaussian noise contribution at the output of the transformation dominates the multiple access interference. In both cases, the decision statistic appears asymptotically Gaussian. The accuracy of the MMSE Gaussian approximation method is verified by several analytical results in [59]. Figure 4.4 depicts the accuracy of the Gaussian approximation BEP for the MMSE detector when compared with the exact calculated BEP. In [59] it also showed that the MMSE BEP is upper bounded by the decorrelating detector BEP.

Another expression of the Gaussian approximated BEP in terms of $J_{\min}$ is given in [7] as

$$P_e(\sigma,1) \approx Q\left(\sqrt{\frac{1-J_{\min}}{J_{\min}}}\right). \tag{4.81}$$

A further approximation of (4.81) is given by

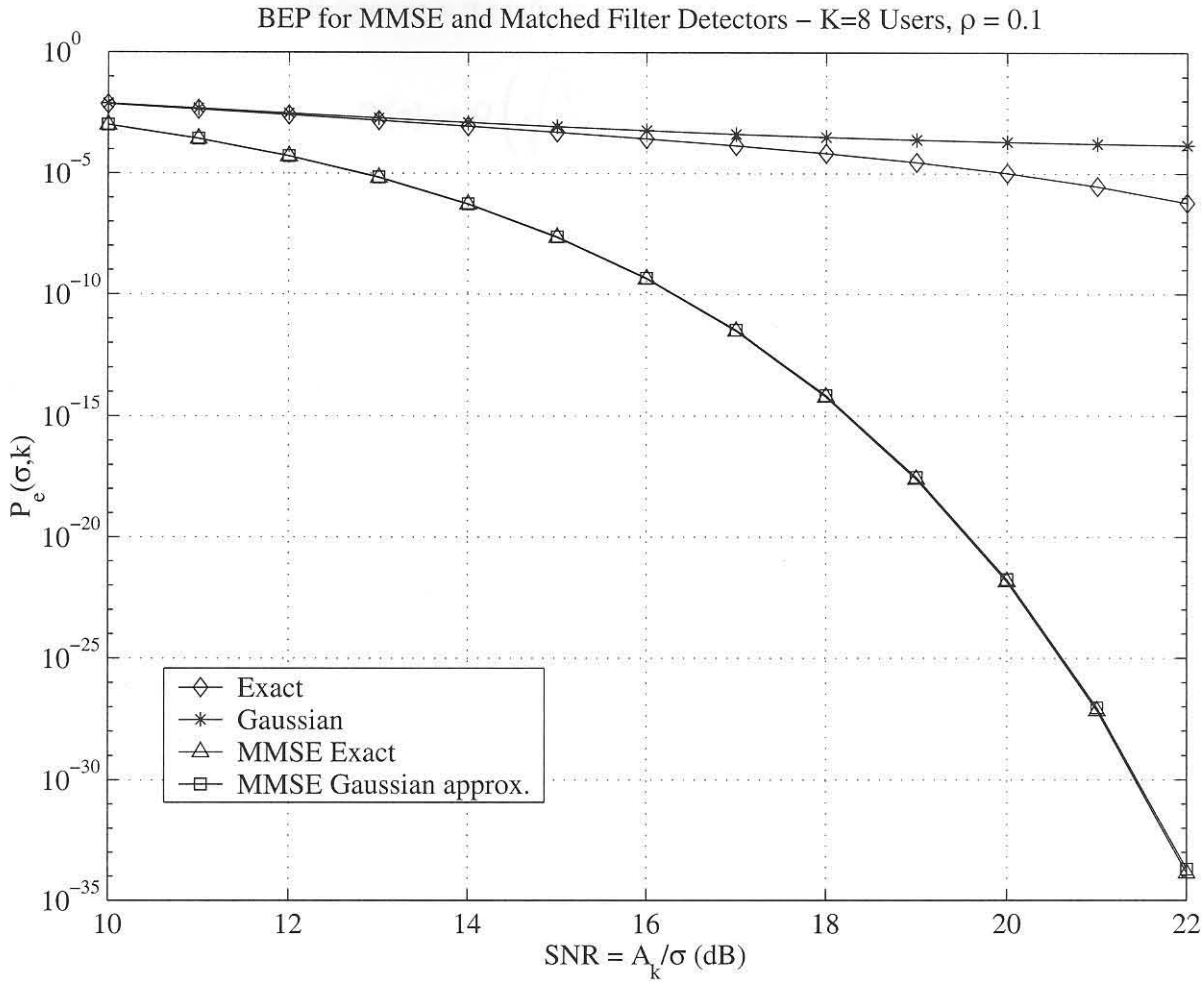

Figure 4.4: BEP graph comparing the exact and Gaussian approximated curves of the MF and MMSE detectors.

$$P_e(\sigma, 1) \approx Q\left(\frac{1}{\sqrt{J_{\min}}}\right). \tag{4.82}$$

4.5.3.2 INFINITE USER LIMIT OF THE MMSE DETECTOR BEP

The infinite user limit BEP is of interest when we consider averaging over random binary sequences. The derivation is rather involved, and we will supply only the result as stated in [31]. It is assumed that all the users have equal power. If the ratio of the number of users to the spreading gain is, or converges to, a constant

$$\beta = \lim_{K\to\infty} \frac{K}{N}, \quad \beta \in (0, +\infty), \tag{4.83}$$

then the BEP of the MMSE detector in the infinite user limit ($K \to \infty$) is given by

$$P_e(\sigma) \to Q\left(\sqrt{\frac{A^2}{\sigma^2} - \frac{1}{4}\mathcal{F}\left(\frac{A^2}{\sigma^2}, \beta\right)}\right), \tag{4.84}$$

where

$$\mathcal{F}(x,z) \stackrel{\text{def}}{=} \left(\sqrt{x\left(1+\sqrt{z}\right)^2+1} - \sqrt{x\left(1-\sqrt{z}\right)^2+1}\right)^2. \tag{4.85}$$

4.5.4 POWER TRADEOFF REGIONS OF THE MMSE DETECTOR

Using the results in the previous sections, we can now determine the power tradeoff regions of the MMSE (and related blind) detectors in the real two user scenario. In Figure 4.5 it can be seen that for all but very high cross correlation values ρ, the SNR needed to attain a BEP of less than 3×10^{-5} for both users is slightly above 12dB for a two user system. If we compare this to the matched filter case in Figure 3.6, we find that the SNR needed does not increase along with the interfering user's amplitude. This means that the MMSE detector is effective in mitigating the near far problem, and the interferer's power has no effect on the desired user's bit rate.

4.5.5 MMSE DETECTOR PERFORMANCE IN MULTIPATH CHANNELS

Having looked at the performance of the MMSE detector in synchronous (non-multipath) channels, we will now consider how the detector operates in multipath channels. An extensive evaluation of the performance of the MMSE detector in a multipath environment was done in [47]. We will follow a similar approach using an asynchronous version of our orthonormal projection model in (2.25). We are interested in the performance of the MMSE detector both in terms of minimum mean-square error and BEP. Concerning our derivation, the following important assumptions are made:

1. The received signal window length is equal to one symbol period;
2. No multipath component is later than one symbol period;
3. The receiver is synchronized to the first multipath component;

With this in mind, we can visualize the multipaths of user k as depicted in Figure 4.6.
Within the received signal window, any two multipath components have a correlated part due to the present bit, and an uncorrelated part due to the preceding bit of the later path. The correlated part can be seen as part of the desired signal and a useful diversity component. The uncorrelated part belongs to the preceding bit, which can be viewed as interference. Using our existing model, we will now derive the MMSE in the case of multipath.

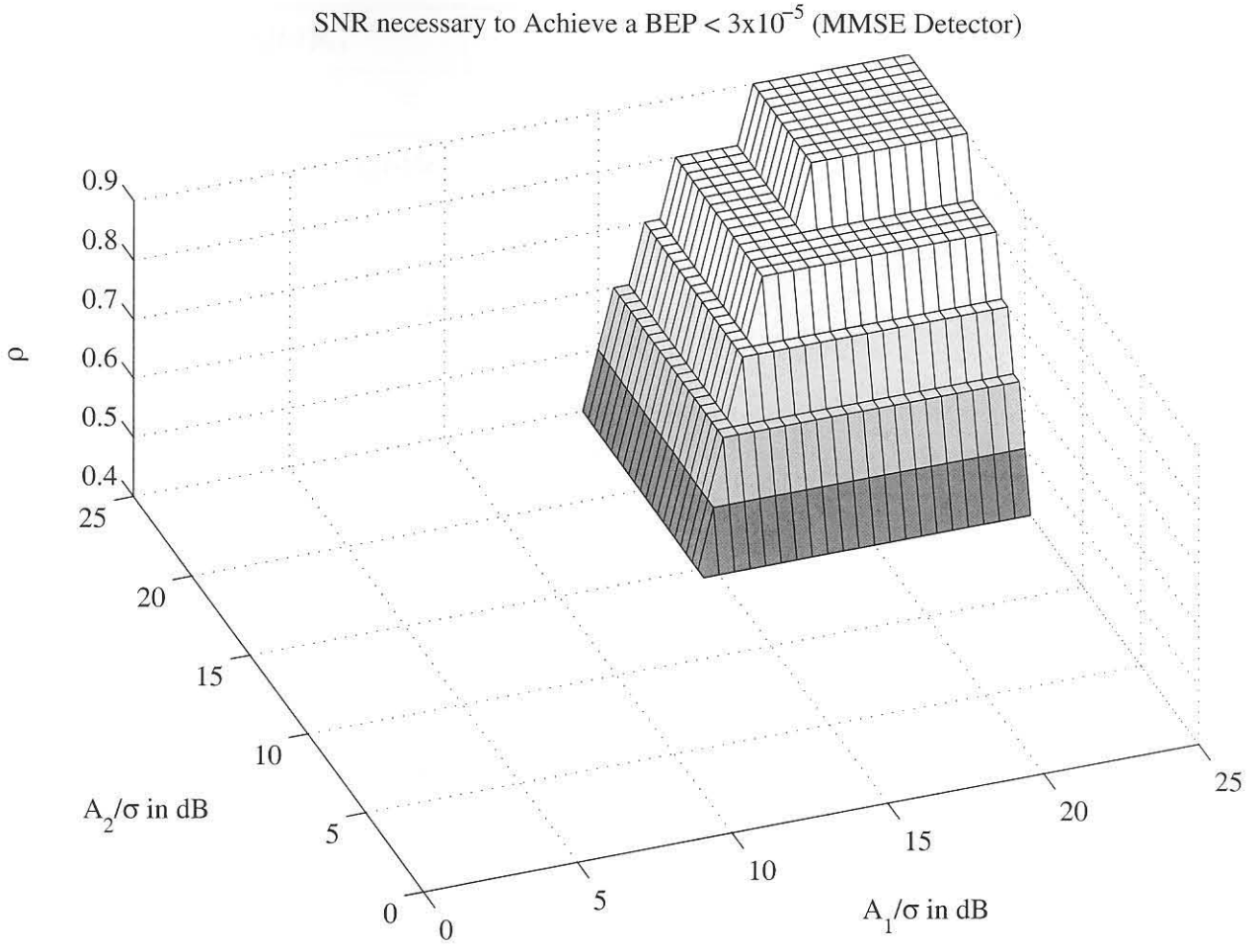

Figure 4.5: Regions of signal-to-noise ratios to attain a BEP of 3×10^{-5} for both users using a MMSE detector.

Remember from (2.59) that when the received signal window is one symbol long, we have

$$y(t)[i] = \sum_{k=1}^{K} \sum_{p=0}^{P-1} A_{k,p} b_k[i] s(\tau - \tau_p) \exp(-j\theta_{k,p}) + \sigma n(t). \tag{4.86}$$

Formulating an asynchronous version of equation (2.25), we have

$$\mathbf{r} = \sum_{k=1}^{K} \sum_{p=0}^{P-1} \left(\tilde{A}_{k,p} b_k[i] \mathbf{s}_{k,p}^{R} + \tilde{A}_{k,p} b_k[i-1] \mathbf{s}_{k,p}^{L} \right) + \sigma \mathbf{m}, \tag{4.87}$$

where the term $\tilde{A}_{k,p}$ refers to the complex amplitude due to the phase term $\theta_{k,p}$ in (4.86), P is the number of resolvable multipaths, and

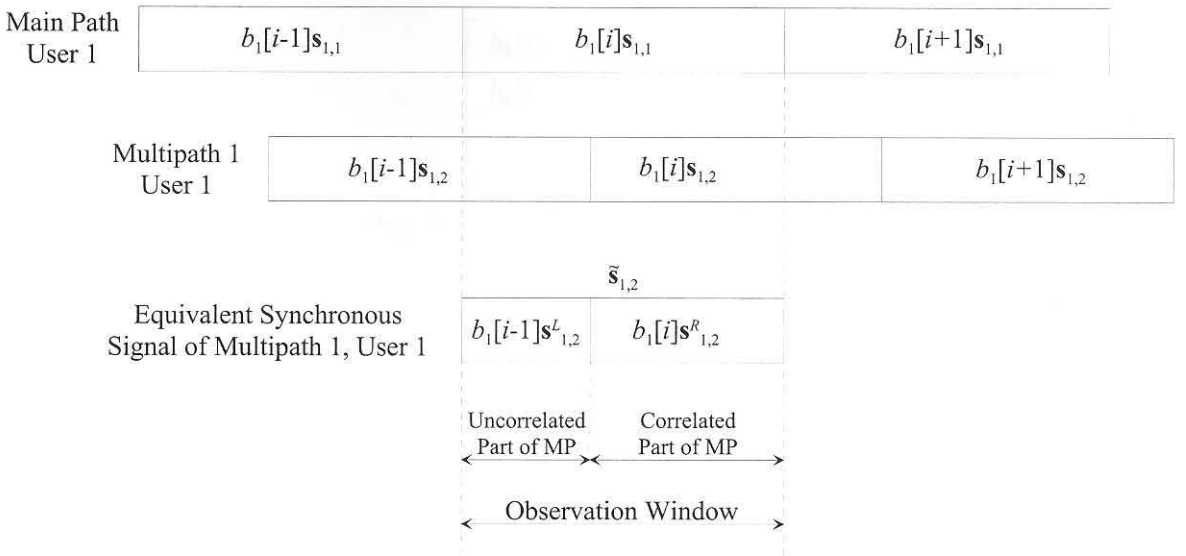

Figure 4.6: Depiction of the equivalent synchronous multipath model of a CDMA channel.

$$\mathbf{s}^L_{k,p} = \begin{bmatrix} s_{k(L-D_p+1)} \\ s_{k(L-D_p+2)} \\ \vdots \\ s_{kL} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{4.88}$$

$$\mathbf{s}^R_{k,p} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ s_{k1} \\ s_{k2} \\ \vdots \\ s_{k(L-D_p)} \end{bmatrix}. \tag{4.89}$$

The symbol s_{kl} defined in (2.23) denotes the the projection of the lth orthonormal signal on the signature waveform of user k. The symbol D_p denotes the delay of the pth multipath. We can let

$$\tilde{\mathbf{s}}_{k,p} = \begin{bmatrix} b_k[i-1]s_{k(L-D_p+1)} \\ b_k[i-1]s_{k(L-D_p+2)} \\ \vdots \\ b_k[i-1]s_{kL} \\ b_k[i]s_{k1} \\ b_k[i]s_{k2} \\ \vdots \\ b_k[i]s_{k(L-D_p)} \end{bmatrix}, \tag{4.90}$$

then

$$\mathbf{r} = \sum_{k=1}^{K} \sum_{p=0}^{P-1} \tilde{A}_{k,p} \tilde{\mathbf{s}}_{k,p} + \sigma \mathbf{m}. \tag{4.91}$$

Analogous to (2.27) we have the covariance matrix

$$E\left[\mathbf{r}\mathbf{r}^H\right] = 2\sigma^2\mathbf{I} + E\left[2\sum_{k=1}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\tilde{\mathbf{s}}_{k,p}\tilde{\mathbf{s}}_{k,p}^H\right] \tag{4.92}$$

$$= E\left[2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\tilde{\mathbf{s}}_{1,p}\tilde{\mathbf{s}}_{1,p}^H + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\tilde{\mathbf{s}}_{k,p}\tilde{\mathbf{s}}_{k,p}^H\right] + 2\sigma^2\mathbf{I}$$

$$= E\left[2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\left(b_1[i-1]\mathbf{s}_{1,p}^L + b_1[i]\mathbf{s}_{1,p}^R\right)\left(b_1[i-1]\mathbf{s}_{1,p}^{L\,H} + b_1[i]\mathbf{s}_{1,p}^{R\,H}\right)\right.$$
$$\left. +2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\left(b_k[i-1]\mathbf{s}_{k,p}^L + b_k[i]\mathbf{s}_{k,p}^R\right)\left(b_k[i-1]\mathbf{s}_{k,p}^{L\,H} + b_k[i]\mathbf{s}_{k,p}^{R\,H}\right)\right] + 2\sigma^2\mathbf{I}$$

$$= 2\sigma^2\mathbf{I} + 2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{R\,H} + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{L\,H}$$
$$+E\left[2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 b_1[i-1]b_1[i]\mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{R\,H}\right] \tag{4.93}$$
$$+E\left[2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 b_1[i]b_1[i-1]\mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{L\,H}\right] \tag{4.94}$$
$$+2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{L\,H} + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{R\,H}$$
$$+E\left[2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 b_k[i-1]b_k[i]\mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{R\,H}\right] \tag{4.95}$$
$$+E\left[2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 b_k[i]b_k[i-1]\mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{L\,H}\right] \tag{4.96}$$

$$= \underbrace{2\sigma^2\mathbf{I}}_{A} + \underbrace{2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H}_{B} + \underbrace{2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{R\,H}}_{C} + \underbrace{2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{L\,H}}_{D}$$
$$\underbrace{+2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{L\,H}}_{E} + \underbrace{2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{R\,H}}_{F}. \tag{4.97}$$

Note that we retain the expected value in (4.92), since the dependence on the current and previous bit is contained in $\tilde{\mathbf{s}}_{k,p}$. When we expand (4.92), the terms (4.93), (4.94), (4.95) and (4.96) become zero, as consecutive bits are uncorrelated. The terms in (4.97) require some further explanation. The term A denotes the sum of the AWGN due to all the multipaths. The terms B and C denote the contributions of the synchronous first multipath and the correlated parts of the other paths respectively of user 1.

These are the signals of interest. Term D denotes the uncorrelated parts of the other multipaths of user 1. The term E denotes the correlated parts of the multipaths of the remaining users, while term F denotes the uncorrelated parts of the multipaths of the remaining users. The terms E and F can both be seen as the contribution of multiple access interference.

The cross correlation vector $\mathbf{p}$, following the same reasoning as in (4.43), is given by the correlation between the vector $\mathbf{r}$ and desired response b_1, i.e

$$\begin{aligned}
\mathbf{p} &= E\left[b_1^*[i]\mathbf{r}\right] \\
&= E\left[b_1^*[i]\left(\sum_{k=1}^{K}\sum_{p=0}^{P-1}\left(\tilde{A}_{k,p}b_k[i]\mathbf{s}_{k,p}^R + \tilde{A}_{k,p}b_k[i-1]\mathbf{s}_{k,p}^L\right) + \sigma\mathbf{m}\right)\right] \\
&= E\left[\sum_{k=1}^{K}\sum_{p=0}^{P-1}\tilde{A}_{k,p}b_k[i]b_1^*[i]\mathbf{s}_{k,p}^R\right] + E\left[\sum_{k=1}^{K}\sum_{p=0}^{P-1}\tilde{A}_{k,p}b_k[i-1]b_1^*[i]\mathbf{s}_{k,p}^L\right] + E\left[b_1^*[i]\sigma\mathbf{m}\right] \\
&= 2\tilde{A}_{1,1}\mathbf{s}_1 + 2\sum_{p=1}^{P-1}\tilde{A}_{1,p}\mathbf{s}_{1,p}^R, \qquad (4.98)
\end{aligned}$$

where (4.98) follows from the fact that the ith bit of user 1 is uncorrelated with the bits of the other users, the previous bits of user 1 and the other users, and the AWGN. If we let

$$\mathbf{g}_1 = \mathbf{s}_1 + \sum_{p=1}^{P-1}\frac{\tilde{A}_{1,p}}{\tilde{A}_{1,1}}\mathbf{s}_{1,p}^R \qquad (4.99)$$

then

$$\mathbf{p} = 2\tilde{A}_{1,1}\mathbf{g}_1. \qquad (4.100)$$

Calculating the optimum solution for the vector transform $\bar{\mathbf{v}}$ for the MMSE multipath case, similar to (2.45), we have

$$\begin{aligned}
\bar{\mathbf{v}} &= \left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1}E\left[b_1^*\mathbf{r}\right] \\
&= 2\tilde{A}_{1,1}\left[2\sigma^2\mathbf{I} + 2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^R{\mathbf{s}_{1,p}^R}^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^L{\mathbf{s}_{1,p}^L}^H\right. \\
&\quad \left. + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^L{\mathbf{s}_{k,p}^L}^H + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^R{\mathbf{s}_{k,p}^R}^H\right]^{-1}\mathbf{g}_1 \qquad (4.101)
\end{aligned}$$

Similar to the non-multipath case, we express the minimum mean-square error as

$$
\begin{aligned}
J_{\min} &= 2 - \mathbf{p}^H \mathbf{C}^{-1} \mathbf{p} \\
&= 2 - 2\left|\tilde{A}_{1,1}\right|^2 \mathbf{g}_1^H \left[\sigma^2 \mathbf{I} + \left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1 \mathbf{s}_1^H + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^R \mathbf{s}_{1,p}^{R\,H} + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L \mathbf{s}_{1,p}^{L\,H}\right. \\
&\quad \left. + \sum_{k=2}^{K} \sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L \mathbf{s}_{k,p}^{L\,H} + \sum_{k=2}^{K} \sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R \mathbf{s}_{k,p}^{R\,H}\right]^{-1} \mathbf{g}_1
\end{aligned} \tag{4.102}
$$

4.5.5.1 SIGNAL-TO-INTERFERENCE RATIO OF THE MMSE DETECTOR IN A MULTIPATH CHANNEL

We define the interference covariance matrix in the multipath case as

$$
\tilde{\mathbf{\Omega}} \stackrel{\text{def}}{=} \sigma^2 \mathbf{I} + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L \mathbf{s}_{1,p}^{L\,H} + \sum_{k=2}^{K} \sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L \mathbf{s}_{k,p}^{L\,H} + \sum_{k=2}^{K} \sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R \mathbf{s}_{k,p}^{R\,H}, \tag{4.103}
$$

and the covariance matrix of the desired component as

$$
\tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H \stackrel{\text{def}}{=} \mathbf{s}_1 \mathbf{s}_1^H + \sum_{p=1}^{P-1} \frac{\left|\tilde{A}_{1,p}\right|^2}{\left|\tilde{A}_{1,1}\right|^2} \mathbf{s}_{1,p}^R \mathbf{s}_{1,p}^{R\,H}, \tag{4.104}
$$

where

$$
\tilde{\mathbf{S}}_1 = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_{1,1}^R & \cdots & \mathbf{s}_{1,P-1}^R \end{bmatrix} \tag{4.105}
$$

is a $L \times P$ matrix. Since the product $\tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H$ is not a scalar, we cannot use the same simplification as (4.63) and (4.64) by using (4.65). We can simply write the optimum MMSE transformation in a multipath channel as

$$
\tilde{\mathbf{v}} = 2\tilde{A}_{1,1} \left[\tilde{\mathbf{\Omega}} + \left|\tilde{A}_{1,1}\right|^2 \tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H\right]^{-1} \mathbf{g}_1, \tag{4.106}
$$

and the minimum mean-square error as

$$
J_{\min} = 2 - 2\left|\tilde{A}_{1,1}\right|^2 \mathbf{g}_1^H \left[\tilde{\mathbf{\Omega}} + \left|\tilde{A}_{1,1}\right|^2 \tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H\right]^{-1} \mathbf{g}_1. \tag{4.107}
$$

The SIR of user 1 in the multipath case is given by

$$
\gamma_{c1} = |A_{1,1}|^2 \mathbf{g}_1^H \tilde{\mathbf{\Omega}}^{-1} \mathbf{g}_1, \tag{4.108}
$$

where $\mathbf{g}_1^H \mathbf{g}_1$ is the gain due multipath. The loss due to the uncorrelated part of the multipaths of user 1, as well as the multiple access interference is contained in $\tilde{\mathbf{\Omega}}$.

4.5.5.2 BEP OF THE MMSE DETECTOR IN A MULTIPATH CHANNEL

Evaluating the exact BEP of the MMSE detector in a multipath environment is even more computationally expensive than in the non-multipath case. To evaluate the BEP of the MMSE detector in a multipath environment we will simply use the approximation in (4.81) and (4.82), i.e.

$$P_e(\sigma, 1) \approx Q\left(\sqrt{\frac{1 - J_{\min}}{J_{\min}}}\right), \tag{4.109}$$

and

$$P_e(\sigma, 1) \approx Q\left(\frac{1}{\sqrt{J_{\min}}}\right). \tag{4.110}$$

4.6 SUMMARY

In this chapter a rigorous analysis of the MMSE detector, within the context of linear detectors, is undertaken. The blind detectors explored in this dissertation have the same vector weight solutions as the MMSE detector. This necessitates a thorough understanding of the operation and performance of the MMSE detector.

The linear decorrelating detector is introduced in the first section of this chapter. The linear decorrelating detector bears a close resemblance to the MMSE detector, as it performs the same operation as the MMSE detector in the noise free case. The linear multiuser detection optimization problem is then generalized to the finding of the best linear detector. The K user case does not permit a closed form solution to this optimization problem. The two user case is subsequently examined, which does permit a closed form solution.

Following this general view of linear multiuser detectors, focus is then shifted to the operation of the joint linear MMSE detector. The MMSE optimization problem is presented, and is solved through use of the complex valued MMSE detector vector matrix model. The two user MMSE detector is briefly considered. The noise limiting forms of the MMSE detector is then discussed, with focus on the relation between the linear decorrelating detector and the linear MMSE detector. The asynchronous linear MMSE detector model is briefly presented. The Wiener characterization of the linear MMSE detector is subsequently considered, where optimization is reduced from joint optimization, to optimization with respect to only one of the users. In the ensuing section the LMS algorithm for the linear MMSE detector is derived.

The rest of the chapter focuses on the performance of the linear MMSE detector based on the criteria stated in Chapter 3. The performance criteria considered include SIRs, asymptotic multiuser

efficiency, BEP and power tradeoff regions. In the case of BEP, the Gaussian approximation method and the infinite user limit for the MMSE detector is also considered. The following section contains the extension of the model in Verdu [31] to the complex valued multipath case. Certain assumptions are made, and the expressions for SIR and BEP are derived for the multipath channel.

CHAPTER FIVE

BLIND MULTIUSER DETECTION USING THE CONSTANT MODULUS CRITERION

In this chapter, the focus is narrowed to the field of blind linear multiuser detection employing two different modified versions of the constant modulus criterion. The first section is an introduction, presenting the (rather short) history of blind multiuser detection for the CDMA channel. In the second section, the linearly constrained constant modulus criterion is thoroughly analyzed along with the convexity of the cost function. For the *first time*, a global condition is presented with proof for the convexity of the cost function. The derivation of the linearly constrained constant modulus algorithm is also presented in this section. The third section contains the analysis of the linearly constrained *differential* constant modulus criterion as presented by [60]. The convexity of this cost function is analyzed, and is shown to be globally convex. The linearly constrained differential constant modulus algorithm is also derived and presented in this section. In the final section of this chapter, the performance of the aforementioned criteria is discussed at the hand of the results obtained from this chapter.

5.1 INTRODUCTION

Blind multiuser detection was first conceptualized by Honig et. al. [3], and has been based on the principle of the linearly constrained minimum variance (LCMV), which was originally developed for adaptive array antennas [61]. In [3], Honig describes the blind LCMV detector in terms of a *canonical representation* for the linear detector in the signal space. The principle of the LCMV detector is to minimize the receiver output variance, without cancelling the desired signal component. When a stochastic gradient algorithm is used, the solution of the mean weight vector is equivalent to that of the MMSE solution. The stochastic gradient algorithm in the case of the LCMV receiver is termed the linearly constrained minimum variance algorithm (LCMVA). The LCMV detector has the

disadvantage that it may cancel out the desired signal component at the receiver output if there are inaccuracies in the desired signal signature vector. An accurate signature vector estimation is needed for the linear constraint. Another disadvantage of the LCMV detector, is that the weight vector adjusted by the LCMVA fluctuates around the optimum point [3], so that the BEP performance degrades.

Another blind approach which is often used in multipath equalization is the constant modulus algorithm (CMA) [62], [63]. The CMA cannot be directly applied to the CDMA channel, as the weight vector might converge to one of the interfering user signature vectors rather than the desired user signature vector [64]. To overcome this problem, the linearly constrained CMA (LCCMA) was proposed by Miguez and Castedo in [20]. Corrections to the aforementioned paper was introduced in [21] and an incorrect closed form analysis of the LCCMA was done in [22], which was later corrected in [23]. The principle of the linearly constrained constant modulus (LCCM) detector is to minimize the deviation of the receiver output from a constant modulus without cancelling the desired signal component. This means that the desired signal component can be protected from being significantly cancelled even if there are inaccuracies in the estimate of the desired signal vector [20]. Moreover, when the receiver output approaches the target constant modulus, the variance of the weight vector as adjusted by the LCCMA can be expected to be relatively small. These qualities make the LCCMA superior to the LCMVA; however, it has been shown that the LCCMA cannot converge to the optimal point if the desired user amplitude is less than a critical value [60]. To overcome this problem, Miyajima /citeMiyajima00 proposed the linearly constrained differential constant modulus (LCDCM) detector to negate the limitation on the desired user amplitude. The stochastic gradient algorithm employing the LCDCM criterion, is subsequently called the linearly constrained differential CMA (LCDCMA). In this dissertation it will be shown that the LCDCM detector achieves comparable performance to the LCCM detector, while there is no limitation on the desired user amplitude.

Both the LCMVA and the LCCMA have the disadvantage that in a frequency selective channel, multiple propagation paths are suppressed rather than combined [47]. The author in [47] proposes a multichannel LCCMA (MLCCMA) to perform the task of joint blind multiuser detection and equalization or multipath diversity combination. In [60] it is implied that the multipath channel impulse response can be estimated using a subspace method, and used as the linear constraint for the LCDCMA. However, this approach requires singular value decomposition (SVD) which makes this method computationally expensive.

In this chapter we will thoroughly analyze the LCCM and LCDCM detectors. We will investigate the cost functions of each of the detectors and then derive the stochastic gradient algorithms associated with each cost function.

5.2 THE LINEARLY CONSTRAINED CONSTANT MODULUS CRITERION

The linearly constrained constant modulus (LCCM) cost function is given by

$$\underset{\mathbf{v}^H\mathbf{s}_1=1}{J(\mathbf{v})} = \frac{1}{2}E\left[\left(|y_t|^2-\alpha\right)^2\right], \tag{5.1}$$

subject to the linear constraint $\mathbf{v}^H\mathbf{s}_1 = 1$, where y_t is the transformed received signal and α is an arbitrary real scalar. Since, $y_t = \mathbf{v}^H\mathbf{r}$, the LCCM cost can now be written as

$$\underset{\mathbf{v}^H\mathbf{s}_1=1}{J(\mathbf{v})} = \frac{1}{2}E\left[\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\right] - \alpha E\left[\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\right] + \frac{1}{2}\alpha^2. \tag{5.2}$$

At this point we make a few assumptions concerning our model. First we assume a synchronous channel. Furthermore, we assume that the Gaussian noise component $\sigma \to 0$, and that the signature waveforms are spanned by $\{\psi_1,\ldots,\psi_L\}$. This leaves us with a K dimensional cost function $J(\mathbf{v})$. In the noise free case, $\mathbf{v}^H\mathbf{r} = \mathbf{v}^H\mathbf{S}\mathbf{A}\mathbf{b}$. If we let $u_k = A_k(\mathbf{v}^H\mathbf{s}_k)$ and $\mathbf{u} = [u_1, u_2, \ldots, u_K]^T$, then we can write the cost function $J(\mathbf{v})$ as

$$\underset{u_1=A_1}{J(\mathbf{u})} = \frac{1}{2}E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] - \alpha E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] + \frac{1}{2}\alpha^2, \tag{5.3}$$

where the linear constraint $\mathbf{v}^H\mathbf{s}_1 = 1$ implies that $u_1 = A_1$.

Since $b_k \in \{\pm 1 \pm j\}$ and for different k, b_k are independent random variables, we have the two expected value terms in (5.3) respectively equal to[1]

$$E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] = 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4 \tag{5.4}$$

and

$$E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] = 2\mathbf{u}^H\mathbf{u}. \tag{5.5}$$

Having removed the influence of the expected values on (5.3), we can now write this equation as

$$\underset{u_1=A_1}{J(\mathbf{u})} = 4\left(\mathbf{u}^H\mathbf{u}\right)^2 - 2\sum_{k=1}^{K}|u_k|^4 - 2\alpha\mathbf{u}^H\mathbf{u} + \frac{1}{2}\alpha^2. \tag{5.6}$$

If we write (5.6) in terms of u_k and u_k^*, we have

[1]The two expected value terms are evaluated in Appendix D

$$\underset{u_1=A_1}{J(\mathbf{u})} = 4\left(\sum_{k=1}^{K} u_k^* u_k\right)^2 - 2\sum_{k=1}^{K}(u_k^* u_k)^2 - 2\alpha\sum_{k=1}^{K} u_k^* u_k + \frac{1}{2}\alpha^2. \tag{5.7}$$

Expanding the terms in (5.7) we obtain

$$J(\mathbf{u}) = 4\left(A_1^* A_1 + \sum_{k=2}^{K} u_k^* u_k\right)^2 - 2(A_1^* A_1)^2 - 2\sum_{k=2}^{K}(u_k^* u_k)^2 - 2\alpha A_1^* A_1 - 2\alpha\sum_{k=2}^{K} u_k^* u_k + \frac{1}{2}\alpha^2. \tag{5.8}$$

5.2.1 THE CONVEXITY OF THE LCCM COST FUNCTION

To investigate the convexity of the cost function $J(u_k)$, a property of a continuous convex function in [65] (Theorem 10.2) will be applied. This theorem states that if

$$J\left(\frac{\mathbf{u}_1 + \mathbf{u}_2}{2}\right) \leq \frac{J(\mathbf{u}_1) + J(\mathbf{u}_2)}{2} \tag{5.9}$$

for any points $\mathbf{u}_1$ and $\mathbf{u}_2$, then the function J is *convex*. The function J is *strictly convex* if the above inequality is true as a strict inequality.

Let the projection of $\mathbf{u}$ with $u_1 = 0$ be denoted by $\bar{\mathbf{u}}$. If we write (5.8) in terms of vector norms of $\bar{\mathbf{u}}$, we have

$$J(\mathbf{u}) = 4\left(A_1^* A_1 + \|\bar{\mathbf{u}}\|^2\right)^2 - 2(A_1^* A_1)^2 - 2\|\bar{\mathbf{u}}\|^4 - 2\alpha A_1^* A_1 - 2\alpha\|\bar{\mathbf{u}}\|^2 + \frac{1}{2}\alpha^2. \tag{5.10}$$

Let us start with the RHS of the inequality in equation (5.9),

$$\begin{aligned}
\frac{J(\mathbf{u}_1)}{2} &= 2\left(A_1^* A_1 + \|\bar{\mathbf{u}}_1\|^2\right)^2 - (A_1^* A_1)^2 - \|\bar{\mathbf{u}}_1\|^4 - \alpha A_1^* A_1 - \alpha\|\bar{\mathbf{u}}_1\|^2 + \frac{1}{4}\alpha^2 \\
&= 2(A_1^* A_1)^2 + 2\|\bar{\mathbf{u}}_1\|^4 + 4A_1^* A_1\|\bar{\mathbf{u}}_1\|^2 - (A_1^* A_1)^2 - \|\bar{\mathbf{u}}_1\|^4 - \alpha A_1^* A_1 - \alpha\|\bar{\mathbf{u}}_1\|^2 + \frac{1}{4}\alpha^2 \\
&= \underbrace{(A_1^* A_1)((A_1^* A_1) - \alpha) + \frac{\alpha^2}{4}}_{A} + \underbrace{(4A_1^* A_1 - \alpha)}_{B}\|\bar{\mathbf{u}}_1\|^2 + \|\bar{\mathbf{u}}_1\|^4 \\
&= A + B\|\bar{\mathbf{u}}_1\|^2 + \|\bar{\mathbf{u}}_1\|^4
\end{aligned} \tag{5.11}$$

and equivalently

$$\frac{J(\mathbf{u}_2)}{2} = A + B\|\bar{\mathbf{u}}_2\|^2 + \|\bar{\mathbf{u}}_2\|^4. \tag{5.12}$$

Thus we have the RHS of (5.9) equal to

$$\frac{J(\mathbf{u}_1)+J(\mathbf{u}_2)}{2} = 2A + B\left(\|\bar{\mathbf{u}}_1\|^2+\|\bar{\mathbf{u}}_2\|^2\right)+\|\bar{\mathbf{u}}_1\|^4+\|\bar{\mathbf{u}}_2\|^4. \tag{5.13}$$

The LHS of (5.9) is given by

$$\begin{aligned}
J\left(\frac{\mathbf{u}_1+\mathbf{u}_2}{2}\right) &= 4\left(A_1^*A_1+\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^2}{4}\right)^2-2\left(A_1^*A_1\right)^2-2\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^4}{16}-2\alpha A_1^*A_1\\
&\quad -2\alpha\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^2}{4}+\frac{1}{2}\alpha^2\\
&= 4\left(A_1^*A_1\right)^2+\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^4}{4}+2A_1^*A_1\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^2-2\left(A_1^*A_1\right)^2\\
&\quad -\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^4}{8}-2\alpha A_1^*A_1-\alpha\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^2}{2}+\frac{1}{2}\alpha^2\\
&= 2\left(A_1^*A_1\right)\left(\left(A_1^*A_1\right)-\alpha\right)+\frac{\alpha^2}{2}+\left(2A_1^*A_1-\frac{\alpha}{2}\right)\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^2\\
&\quad +\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|^4}{8}\\
&= 2\left(A_1^*A_1\right)\left(\left(A_1^*A_1\right)-\alpha\right)+\frac{\alpha^2}{2}+4\left(2A_1^*A_1-\frac{\alpha}{2}\right)\left(\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|}{2}\right)^2\\
&\quad +2\left(\frac{\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\|}{2}\right)^4
\end{aligned} \tag{5.14}$$

If we use the triangle inequality $\|\bar{\mathbf{u}}_1+\bar{\mathbf{u}}_2\| \le \|\bar{\mathbf{u}}_1\|+\|\bar{\mathbf{u}}_2\|$ we have

$$\begin{aligned}
J\left(\frac{\mathbf{u}_1+\mathbf{u}_2}{2}\right) &\le 2\left(A_1^*A_1\right)\left(\left(A_1^*A_1\right)-\alpha\right)+\frac{\alpha^2}{2}\\
&\quad +4\left(2A_1^*A_1-\frac{\alpha}{2}\right)\left(\frac{\|\bar{\mathbf{u}}_1\|+\|\bar{\mathbf{u}}_2\|}{2}\right)^2\\
&\quad +2\left(\frac{\|\bar{\mathbf{u}}_1\|+\|\bar{\mathbf{u}}_2\|}{2}\right)^4
\end{aligned} \tag{5.15}$$

Also using the fact that the functions $(\cdot)^2$ and $(\cdot)^4$ are *strictly convex*, we have by (5.9)

$$\begin{aligned}
J\left(\frac{\mathbf{u}_1+\mathbf{u}_2}{2}\right) &< 2\left(A_1^*A_1\right)\left(\left(A_1^*A_1\right)-\alpha\right)+\frac{\alpha^2}{2}\\
&\quad +2\left(2A_1^*A_1-\frac{\alpha}{2}\right)\left(\|\bar{\mathbf{u}}_1\|^2+\|\bar{\mathbf{u}}_2\|^2\right)\\
&\quad +\|\bar{\mathbf{u}}_1\|^4+\|\bar{\mathbf{u}}_2\|^4
\end{aligned} \tag{5.16}$$

If we write (5.16) in terms of A and B, we have

$$\begin{aligned}
J\left(\frac{\mathbf{u}_1+\mathbf{u}_2}{2}\right) &< 2A+B\left(\|\bar{\mathbf{u}}_1\|^2+\|\bar{\mathbf{u}}_2\|^2\right)+\|\bar{\mathbf{u}}_1\|^4+\|\bar{\mathbf{u}}_2\|^4\\
&= \frac{J(\mathbf{u}_1)+J(\mathbf{u}_2)}{2}
\end{aligned} \tag{5.17}$$

Because of the above inequality, and by (5.9), we have proved the *strict convexity* of the LCCM cost function subject to

$$A \geq 0 \tag{5.18}$$

and

$$B \geq 0. \tag{5.19}$$

Since (5.19) is a stricter condition than (5.18), we can discard (5.18). We have thus that the LCCM cost function is convex, subject only to

$$B \geq 0 \text{ or equivalently } A_1^* A_1 \geq \frac{\alpha}{4}. \tag{5.20}$$

Note that the inequality $B > 0$ is a *global condition* on A_1 and α, insuring convexity of the LCCM cost function.

5.2.2 THE STATIONARY POINTS OF THE LCCM COST FUNCTION

Considering equation (5.8) again, if we let $u_k = x_k + jy_k$ for $2 \leq k \leq K$, then the cost function $J(\mathbf{u})$ becomes

$$\begin{aligned} J(\mathbf{u}) &= 4\left(A_1^* A_1 + \sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)\right)^2 - 2\left(A_1^* A_1\right)^2 - 2\sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)^2 \\ &\quad -2\alpha A_1^* A_1 - 2\alpha\sum_{k=2}^{K}\left(x_k^2 + y_k^2\right) + \frac{1}{2}\alpha^2 \end{aligned} \tag{5.21}$$

To solve for the stationary points, we find the gradient (directional derivative) of $J(x_k, y_k)$, and equate it to zero. In this way we can attempt to solve for the points at which the cost function is a minimum. In this case it is more informative to differentiate with respect to the real and imaginary parts of $J(x_k, y_k)$, rather than differentiate with respect to a complex vector:

$$\begin{aligned} \nabla_{x_l} J = \frac{\partial J(\mathbf{u})}{\partial x_l} &= 16\left(A_1^* A_1 + \sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)\right)x_l - 8\left(x_l^2 + y_l^2\right)x_l - 4\alpha x_l \\ &= 4x_l\left(4A_1^* A_1 + 4\sum_{k=2}^{K}\left(x_k^2 + y_k^2\right) - 2\left(x_l^2 + y_l^2\right) - \alpha\right) \\ &= 4x_l\left(4A_1^* A_1 + 4\sum_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right) + 2x_l^2 + 2y_l^2 - \alpha\right) \end{aligned} \tag{5.22}$$

and equivalently

$$\nabla_{y_l} J = \frac{\partial J(\mathbf{u})}{\partial y_l} = 4y_l \left(4A_1^* A_1 + 4 \sum_{\substack{k=2 \\ k \neq l}}^{K} \left(x_k^2 + y_k^2 \right) + 2x_l^2 + 2y_l^2 - \alpha \right) \tag{5.23}$$

with symmetry evident between (5.20) and (5.21). Letting

$$X = \left(4A_1^* A_1 - \alpha + 4 \sum_{\substack{k=2 \\ k \neq l}}^{K} \left(x_k^2 + y_k^2 \right) + 2y_l^2 \right) \tag{5.24}$$

and

$$Y = \left(4A_1^* A_1 - \alpha + 4 \sum_{\substack{k=2 \\ k \neq l}}^{K} \left(x_k^2 + y_k^2 \right) + 2x_l^2 \right), \tag{5.25}$$

we have

$$x_l \left(2x_l^2 + X \right) = 0 \tag{5.26}$$

and

$$y_l \left(2y_l^2 + Y \right) = 0. \tag{5.27}$$

At this stage two cases can be identified:

1. $4A_1^* A_1 - \alpha \geq 0$;
2. $4A_1^* A_1 - \alpha < 0$.

In the case of $4A_1^* A_1 - \alpha \geq 0$, it is evident that a unique solution exists at $x_l = 0$ and $y_l = 0$, since all the other terms in (5.24) and (5.25) can only be greater than or equal to zero. In the case of $4A_1^* A_1 - \alpha < 0$, solutions may exist at $x_l = 0$ or $x_l = \pm\sqrt{\frac{-X}{2}}$ and $y_l = 0$ or $y_l = \pm\sqrt{\frac{-Y}{2}}$.

We have already ascertained that the LCCM cost function is strictly convex for certain values of α and A_1. Let us consider the trivial solution of $x_l = 0, y_l = 0$. This implies that $u_k = 0 + j0$ with $2 \leq k \leq K$ or $\bar{\mathbf{u}} = \mathbf{0}$, where $\bar{\mathbf{u}} = [0 \; u_2 \; u_3 \; \ldots \; u_K]^T$. Coincidentally, this is also the solution that cancels out all multiuser interference. To prove that the point $\bar{\mathbf{u}} = \mathbf{0}$ is a global minimum of the cost function J conditioned on A_1 and α, we will have to look at the Hessian matrix $\mathbf{H}(J)$ of the cost function J at the point $\bar{\mathbf{u}} = \mathbf{0}$.

To evaluate the nature of the stationary point $\bar{\mathbf{u}} = \mathbf{0}$ of the LCCM cost function, we use the Hessian matrix as defined below

$$\mathbf{H}(J) = \begin{bmatrix} \mathbf{H}_a & \mathbf{H}_b \\ \mathbf{H}_c & \mathbf{H}_d \end{bmatrix}, \tag{5.28}$$

with

$$(\mathbf{H}_a)_{ml} = \frac{\partial}{\partial x_m}\left(\frac{\partial J}{\partial x_l}\right), \tag{5.29}$$

$$(\mathbf{H}_b)_{ml} = \frac{\partial}{\partial y_m}\left(\frac{\partial J}{\partial x_l}\right), \tag{5.30}$$

$$(\mathbf{H}_c)_{ml} = \frac{\partial}{\partial x_m}\left(\frac{\partial J}{\partial y_l}\right), \tag{5.31}$$

and

$$(\mathbf{H}_d)_{ml} = \frac{\partial}{\partial y_m}\left(\frac{\partial J}{\partial y_l}\right). \tag{5.32}$$

The entries of the Hessian matrix are

$$(\mathbf{H}_a)_{ml} = \begin{cases} 16A_1^*A_1 + 24x_l^2 + 8y_l^2 + 16\sum\limits_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right) - \alpha & \text{if } l = m \\ 32x_m x_l & \text{if } l \neq m \end{cases}, \tag{5.33}$$

$$(\mathbf{H}_b)_{ml} = 32y_m x_l, \tag{5.34}$$

$$(\mathbf{H}_c)_{ml} = 32x_m y_l, \tag{5.35}$$

and

$$(\mathbf{H}_d)_{ml} = \begin{cases} 16A_1^*A_1 + 24y_l^2 + 8x_l^2 + 16\sum\limits_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right) - \alpha & \text{if } l = m \\ 32x_m x_l & \text{if } l \neq m \end{cases}. \tag{5.36}$$

Normally, for the Hessian matrix to be positive definite (i.e., the cost function convex with a global minimum), the determinants of all the principle submatrices of the Hessian matrix must be zero. This is to say that:

$$\begin{aligned} \det\left((\mathbf{H}_a)_{11}\right) &= (\mathbf{H}_a)_{11} \geq 0, \\ (\mathbf{H}_a)_{11}(\mathbf{H}_a)_{22} - (\mathbf{H}_a)_{12}(\mathbf{H}_a)_{21} &\geq 0, \\ &\vdots \\ \det \mathbf{H}_a &\geq 0, \\ &\vdots \\ \det \mathbf{H}(J) &\geq 0, \end{aligned} \tag{5.37}$$

Since we have already proved that the LCCM cost function is strictly convex subject to (5.20), we can now show that the point $\bar{\mathbf{u}} = \mathbf{0}$ is a unique global minimum subject to the same conditions. Since it is possible that a strictly convex function may have only one minimum, we can prove the point $\bar{\mathbf{u}} = \mathbf{0}$ a global minimum, by proving it a local minimum [66].

Let us now prove that a *local minimum* exists only at the point $\bar{\mathbf{u}} = \mathbf{0}$ for certain values of A_1 and α. Implementing this, we have the diagonal Hessian

$$(\mathbf{H}(J_0))_{ij} = \begin{cases} 16A_1^* A_1 - 4\alpha & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{5.38}$$

For each of the diagonal elements to be ≥ 0, and the matrix *positive semi-definite*, the following must be true:

$$A_1^* A_1 \geq \alpha/4. \tag{5.39}$$

We have thus proved the local minimum of J at $\bar{\mathbf{u}} = \mathbf{0}$ subject to the condition (5.39). Note that the condition in (5.39) is a *local condition* on A_1 and α, and insures only a local minimum. Coincidentally, this condition corresponds to the condition on global convexity in (5.20).

In the preceding text we have proved the global convexity of the LCCM cost function subject to (5.20) by using a definition of a continuous convex function in (5.9). We have also seen that the single stationary point (also subject to (5.20)), is a global minimum due to the convexity of the LCCM cost function. The conditions of $A_1^* A_1 > \alpha/4$, $A_1^* A_1 = \alpha/4$ and $A_1^* A_1 < \alpha/4$ are depicted in Figures 5.1, 5.2 and 5.3 respectively. This clearly supports the notion we have developed regarding the convexity of the LCCM cost function. For the case of $A_1^* A_1 \geq \alpha/4$, the cost function is strictly convex. If $A_1^* A_1 < \alpha/4$, convexity cannot be guaranteed anymore, as can be seen in Figure 5.3. From this figure it is also evident that the point $x = 0$, $y = 0$ is also not the only stationary point. This corresponds to the solutions of equations (5.26) and (5.27).

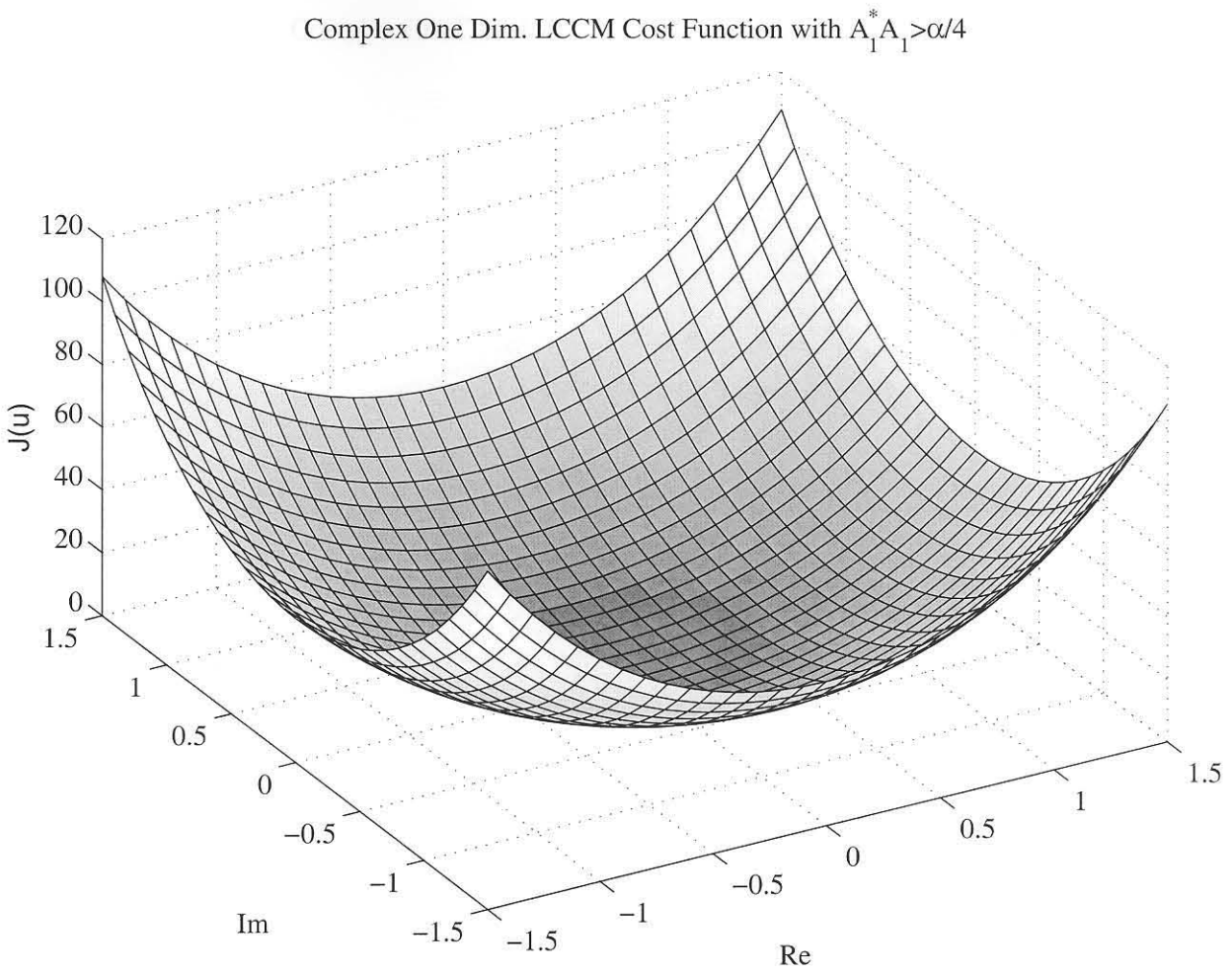

Figure 5.1: Complex LCCM cost function surface with $A_1^* A_1 > \alpha/4$.

5.2.3 LINEARLY CONSTRAINED CONSTANT MODULUS ALGORITHM

The linearly constrained constant modulus algorithm was originally inspired by its application to the field of adaptive arrays [61, 67]. It is based on the generalized sidelobe canceller, which incorporates *a priori* information about the signal. The linear constraint is implemented to capture the user of interest instead of any of the interference signals.

Recall that the LCCM cost function is given by

$$\underset{\mathbf{v}^H \mathbf{s}_1 = 1}{J(\mathbf{v})} = \frac{1}{2} E\left[\left(|y_t|^2 - \alpha\right)^2\right]. \tag{5.40}$$

Let us first consider the unconstrained cost function

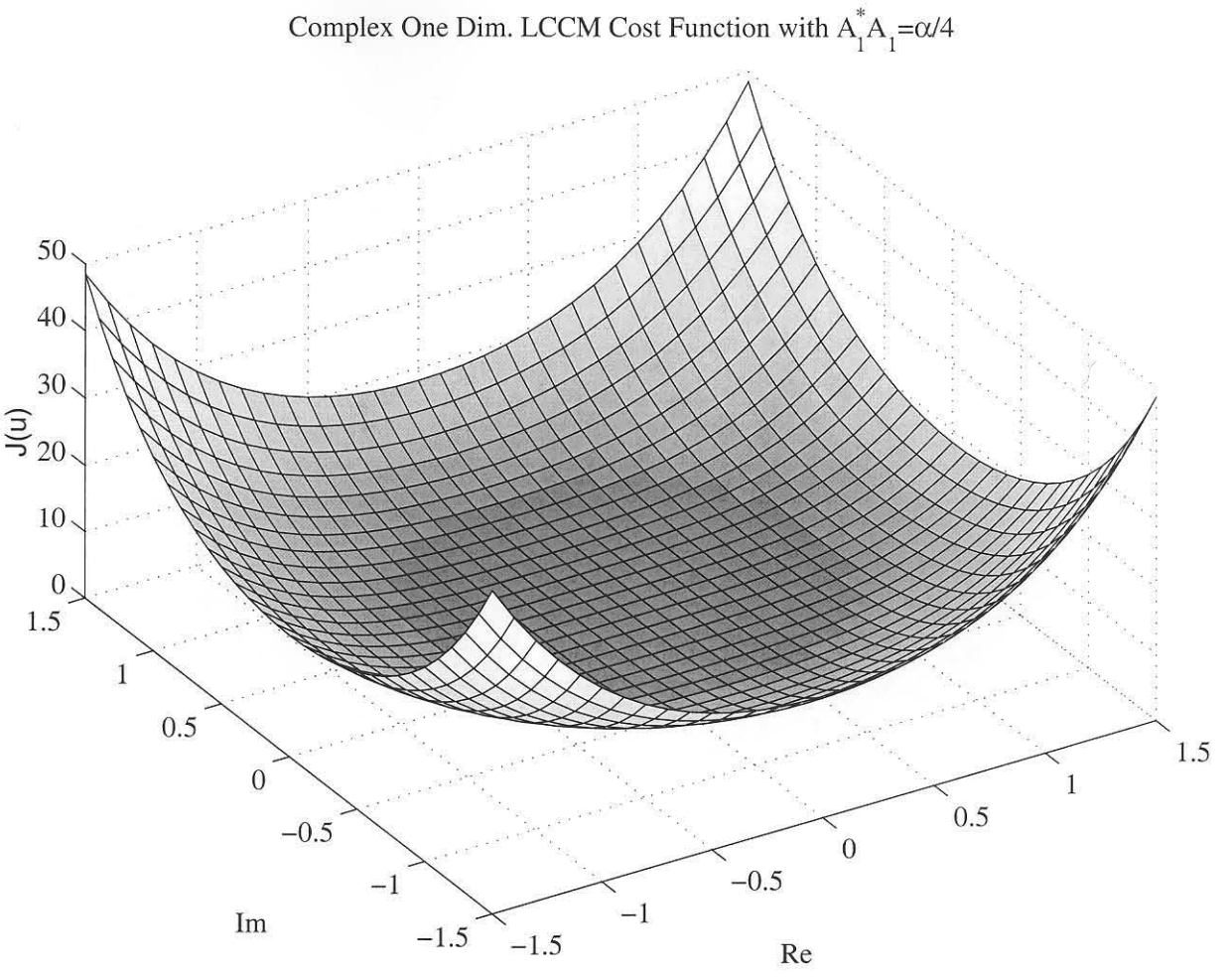

Figure 5.2: Complex LCCM cost function surface with $A_1^* A_1 = \alpha/4$.

$$J = \frac{1}{2} E\left[\left(|y_t|^2 - \alpha\right)^2\right], \tag{5.41}$$

which, if we consider it in terms of inner products, becomes

$$J = \frac{1}{2} E\left[\left(\langle y, c_1^* \rangle^2 - \alpha\right)^2\right], \tag{5.42}$$

where c_1 is (as in the case of the LMS algorithm) the multidimensional parameter which operates on y in the form of a linear transform.

We may consider a *canonical representation* of the linear transform c_1 in terms of the signature waveform of user 1, viz. s_1, and a component orthogonal to s_1, denoted by x_1:

$$c_1 = s_1 + x_1, \tag{5.43}$$

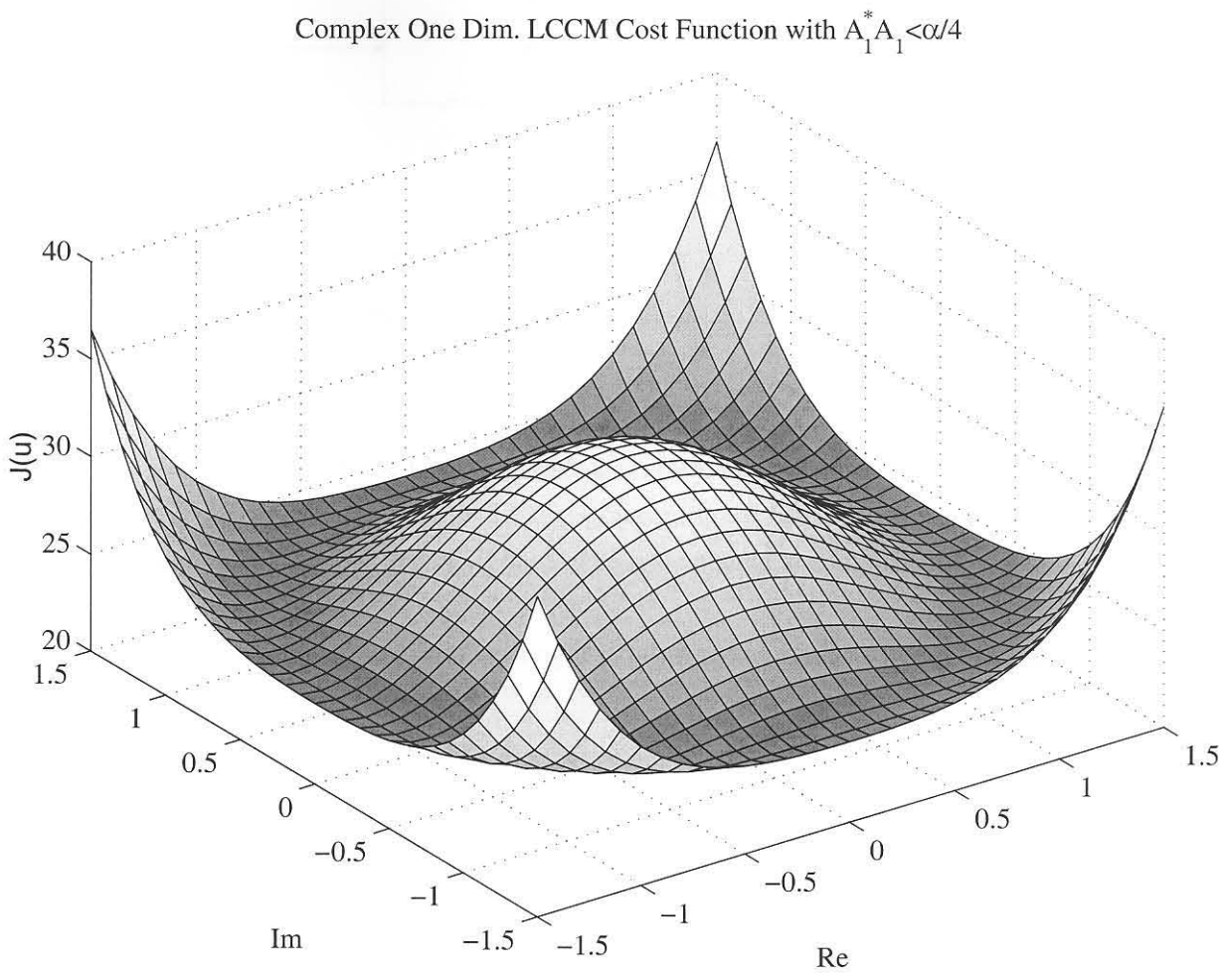

Figure 5.3: Complex LCCM cost function surface with $A_1^* A_1 < \alpha/4$.

where x_1 is such that

$$\langle s_1, x_1^* \rangle = 0 \tag{5.44}$$

This representation is canonical in that every linear multiuser detector of user 1 can be expressed in that form. The set of signals c_1 that can be written as (5.43) and (5.44) are those that satisfy

$$\langle s_1, c_1^* \rangle = \|s_1\|^2 = 1, \tag{5.45}$$

and the decision of $\hat{b}_1 = \text{sgn}(\langle y, c_1^* \rangle)$ is invariant to positive scaling. This means that the only linear transformations that are ruled out by (5.45), are the set of signals c_1 orthogonal to s_1. These signals may be omitted, since they result in an error probability of $1/2$.

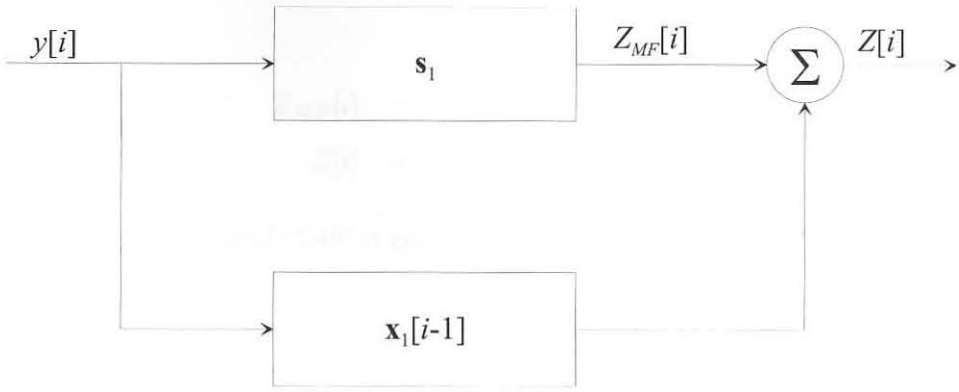

Figure 5.4: Generalized Sidelobe Canceller with $\mathbf{x}[i]$ governed by (5.53) in the case of the LCCM algorithm and (5.79) in the case of the LCDCM algorithm. In the case of the LCDCM algorithm, the previous values $Z_D[i]$, $Z_{MF_D}[i]$ and $\mathbf{r}_D[i]$ need to be remembered to compute $\mathbf{x}[i]$.

Returning to the cost function in (5.43), writing it in terms of the decomposition of c_1 in (5.44) and ignoring the expected value[2], we have

$$J(x_1) = 2\left(\langle y, s_1^* + x_1^* \rangle^2 - \alpha\right)^2. \tag{5.46}$$

If we then find the multidimensional gradient of (5.46), we have

$$\nabla J = 2\left(\langle y, s_1^* + x_1^* \rangle\right)\left(\langle y, s_1^* + x_1^* \rangle^2 - \alpha\right) y^*. \tag{5.47}$$

Note that we are still working with the gradient of the unconstrained cost function. The linear constraint allows the detector to tune out the interference orthogonal to the signature waveform, i.e. restricting the detector from tuning out the desired component. In terms of the gradient, we are looking for the projection or subspace for which the gradient stays orthogonal to s_1. Since the inner product in (5.47) is a complex scaling factor, and y^* is the only multidimensional parameter in the equation, we can restrict y^* (and thus the gradient) to be orthogonal to s_1 by replacing y^* with

$$y - \langle y, s_1^* \rangle s_1. \tag{5.48}$$

Therefore the projection of the gradient in the direction orthogonal to s_1 is

$$\nabla J = 2\left(\langle y, s_1^* + x_1^* \rangle\right)\left(\langle y, s_1^* + x_1^* \rangle^2 - \alpha\right)\left[y - \langle y, s_1^* \rangle s_1\right]. \tag{5.49}$$

Let us denote the matched filter responses for s_1 and $s_1 + x_1[i-1]$ respectively by

[2] As in the case of the LMS algorithm, we may do this. The reason for this is that in the execution of several iterations, the trajectory will be, on average, in the direction of the steepest descent.

$$Z_{MF}[i] = \langle y[i], s_1^* \rangle , \tag{5.50}$$

$$Z[i] = \langle y[i], s_1^* + x_1^*[i-1] \rangle . \tag{5.51}$$

The stochastic adaptation rule of (5.49) is given by

$$x_1[i] = x_1[i-1] - \mu Z[i] \left(Z^2[i] - \alpha \right) \left(y[i] - Z_{MF}[i] s_1 \right) , \tag{5.52}$$

which corresponds to the block diagram of the generalized sidelobe canceller in Figure 5.4. As in the case of the LMS algorithm, we may do the following modifications to our system:

- Implementation with finite dimensional vectors rather than continuous time signals.
- Improved convergence speed with more complex recursive algorithms, such as recursive least squares (RLS).
- Implementation in asynchronous channels.

The finite dimensional vector implementation of our LCCM algorithm is given by

$$\mathbf{x}_1[i] = \mathbf{x}_1[i-1] - \mu Z[i] \left(Z^2[i] - \alpha \right) \left(\mathbf{r}[i] - Z_{MF}[i] \mathbf{s}_1 \right) , \tag{5.53}$$

with

$$Z_{MF}[i] = \mathbf{s}_1^H \mathbf{r}[i], \tag{5.54}$$

$$Z[i] = \left(\mathbf{s}_1 + \mathbf{x}_1[i-1] \right)^H \mathbf{r}[i]. \tag{5.55}$$

5.3 THE LINEARLY CONSTRAINED DIFFERENTIAL CONSTANT MODULUS CRITERION

The linearly constrained differential constant modulus (LCDCM) cost function is given by

$$\underset{\mathbf{v}^H \mathbf{s}_1 = 1}{J(\mathbf{v})} = \frac{1}{2} E \left[\left(|y_t|^2 - |y_{t_D}|^2 \right)^2 \right] , \tag{5.56}$$

subject to the linear constraint $\mathbf{v}^H \mathbf{s}_1 = 1$, where y_t is the transformed received signal and y_{t_D} is a delayed version of the transformed received signal. The LCDCM criterion attempts to keep the modulus of the received signal constant from time t to time $t + D$. Following the same reasoning as in the case of the LCCM detector, we will show that the LCDCM cost function has a *global*

minimum. Since, $y_t = \mathbf{v}^H\mathbf{r}$ and assuming a quasi stationary CDMA channel, the LCDCM cost can now be written as

$$\begin{aligned}\underset{\mathbf{v}^H\mathbf{s}_1=1}{J(\mathbf{v})} &= \tfrac{1}{2}E\left[\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\right] - E\left[\mathbf{v}^H\mathbf{r}\mathbf{r}^H\mathbf{v}\mathbf{v}^H\mathbf{r}_D\mathbf{r}_D^H\mathbf{v}\right] \\ &\quad + \tfrac{1}{2}E\left[\mathbf{v}^H\mathbf{r}_D\mathbf{r}_D^H\mathbf{v}\mathbf{v}^H\mathbf{r}_D\mathbf{r}_D^H\mathbf{v}\right],\end{aligned} \tag{5.57}$$

where r_D is the delayed received vector. Again, we assume that the Gaussian noise component $\sigma \to 0$, and that the signature waveforms are spanned by $\{\psi_1, \ldots, \psi_L\}$. This leaves us with a K dimensional cost function $J(\mathbf{v})$. In the noise free case, $\mathbf{v}^H\mathbf{r} = \mathbf{v}^H\mathbf{S}\mathbf{A}\mathbf{b}$. If we let $u_k = A_k(\mathbf{v}^H\mathbf{s}_k)$ and $\mathbf{u} = [u_1, u_2, \ldots, u_K]^H$, then we can write the cost function $J(\mathbf{v})$ as

$$\begin{aligned}\underset{u_1=A_1}{J(\mathbf{u})} &= \tfrac{1}{2}E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\right] - E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] \\ &\quad + \tfrac{1}{2}E\left[\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right].\end{aligned} \tag{5.58}$$

Let us again assume that the bits of different users are independent and that $b[i], b[i-D] \in \{\pm 1 \pm j\}$. Furthermore, assuming that the delay D is greater than any partial response signalling inherent in the system, we have that bits separated by D seconds are independent, and hence the expectation value terms of (5.58) can be written as

$$E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\right] = 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4, \tag{5.59}$$

$$E\left[\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] = 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4 \tag{5.60}$$

and

$$E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] = 4\left(\mathbf{u}^H\mathbf{u}\right)^2. \tag{5.61}$$

This greatly simplifies the LCDCM cost function to

$$\underset{u_1=A_1}{J(\mathbf{u})} = 4\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4. \tag{5.62}$$

Writing (5.62) in terms of a summation of u_k and exercising the linear constraint $u_1 = A_1$, we have

$$J(\mathbf{u}) = 4\left(A_1^*A_1 + \sum_{k=2}^{K}u_k^*u_k\right)^2 - 4\sum_{k=2}^{K}\left(u_k^*u_k\right)^2 - 4\left(A_1^*A_1\right)^2. \tag{5.63}$$

Letting $u_k = x_k + jy_k$ for $2 \le k \le K$, the cost function $J(\mathbf{u})$ becomes

$$J(\mathbf{u}) = 4\left(A_1^* A_1 + \sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)\right)^2 - 4\sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)^2 - 4\left(A_1^* A_1\right)^2. \tag{5.64}$$

Once again, the gradient of the cost function ∇J with respect to the lth real and imaginary elements of $\mathbf{u}$ is found, and equated it to zero, yields:

$$\begin{aligned}
\nabla_{x_l} J &= 16x_l\left(A_1^* A_1 + \sum_{k=2}^{K}\left(x_k^2 + y_k^2\right)\right) - 16x_l\left(x_l^2 + y_l^2\right) \\
&= 16x_l\left(A_1^* A_1 + x_l^2 + y_l^2 + \sum_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right) - x_l^2 - y_l^2\right) \\
&= 16x_l\left(A_1^* A_1 + \sum_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right)\right) = 0
\end{aligned} \tag{5.65}$$

and equivalently

$$\nabla_{y_l} J = 16y_l\left(A_1^* A_1 + \sum_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right)\right) = 0. \tag{5.66}$$

This is highly encouraging, since there exists a unique trivial solution of $x_l = 0$ and $y_l = 0$ for the gradient equations (5.65) and (5.66). This corresponds to $\bar{\mathbf{u}} = \mathbf{0}$, once again with $\bar{\mathbf{u}} = [u_2\ u_3\ \ldots\ u_K]$, which is also the solution of the MMSE detector with no multipath or code mismatch.

We will now examine if $\bar{\mathbf{u}} = \mathbf{0}$ is a *global minimum*. As will be seen in the following section, we need not even consider the convexity of the function $J(\bar{\mathbf{u}})$ to determine if $\bar{\mathbf{u}} = \mathbf{0}$ is a global minimum.

5.3.1 GLOBAL MINIMUM OF THE LCDCM COST FUNCTION

In the case of the linearly constrained differential CMA, it is simple to show that the cost function has a global minimum, without even having to consider the convexity of the LCDCM cost function. Since the gradient functions (5.65) and (5.66) only has a trivial solution at $\bar{\mathbf{u}} = \mathbf{0}$, we need only to examine the nature of the stationary point $\bar{\mathbf{u}} = \mathbf{0}$. Again we use the Hessian as defined in (5.28). The entries of the Hessian are given by

$$(\mathbf{H}_a)_{ml} = \begin{cases} 16\left(A_1^* A_1 + \sum\limits_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right)\right) & \text{if } l = m \\ 32x_m x_l & \text{if } l \neq m \end{cases}, \tag{5.67}$$

$$(\mathbf{H}_b)_{ml} = 32 y_m x_l, \tag{5.68}$$

$$(\mathbf{H}_c)_{ml} = 32 x_m y_l, \tag{5.69}$$

and

$$(\mathbf{H}_d)_{ml} = \begin{cases} 16\left(A_1^* A_1 + \sum\limits_{\substack{k=2 \\ k\neq l}}^{K}\left(x_k^2 + y_k^2\right)\right) & \text{if } l = m \\ 32 x_m x_l & \text{if } l \neq m \end{cases}. \tag{5.70}$$

The Hessian at the point $x_l = y_l = 0$ is given by

$$\mathbf{H}(J_0) = \begin{bmatrix} 16A_1^* A_1 & 0 & \cdots & 0 \\ 0 & 16A_1^* A_1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 16A_1^* A_1 \end{bmatrix}, \tag{5.71}$$

which is positive definite. This means that the point $x_l = y_l = 0$ or $\bar{\mathbf{u}} = \mathbf{0}$ is a *unique global minimum*.

We have seen in this section that the LCDCM cost criterion exhibits a global minimum. Unlike the LCCM criterion, this point remains a minimum irrespective of desired user amplitude. Figure 5.5 shows the one dimensional complex surface of the LCDCMA cost function for any value of A_1. It is clearly convex with a global minimum.

5.3.2 LINEARLY CONSTRAINED DIFFERENTIAL CONSTANT MODULUS ALGORITHM

The LCDCMA can be derived by also using the stochastic gradient approach as in the case of the LMS and LCCM algorithms. Again, recall the LCDCM cost function:

$$\underset{\mathbf{v}^H \mathbf{s}_1 = 1}{J(\mathbf{v})} = \frac{1}{2} E\left[\left(|y_t|^2 - |y_{t_D}|^2\right)^2\right]. \tag{5.72}$$

Let us first consider the unconstrained cost (as with the LCCMA), which is given by

$$J = \frac{1}{2} E\left[\left(|y_t|^2 - |y_{t_D}|^2\right)^2\right]. \tag{5.73}$$

The cost function in terms of inner products representing the transformed received and delayed transformed received signals is given by

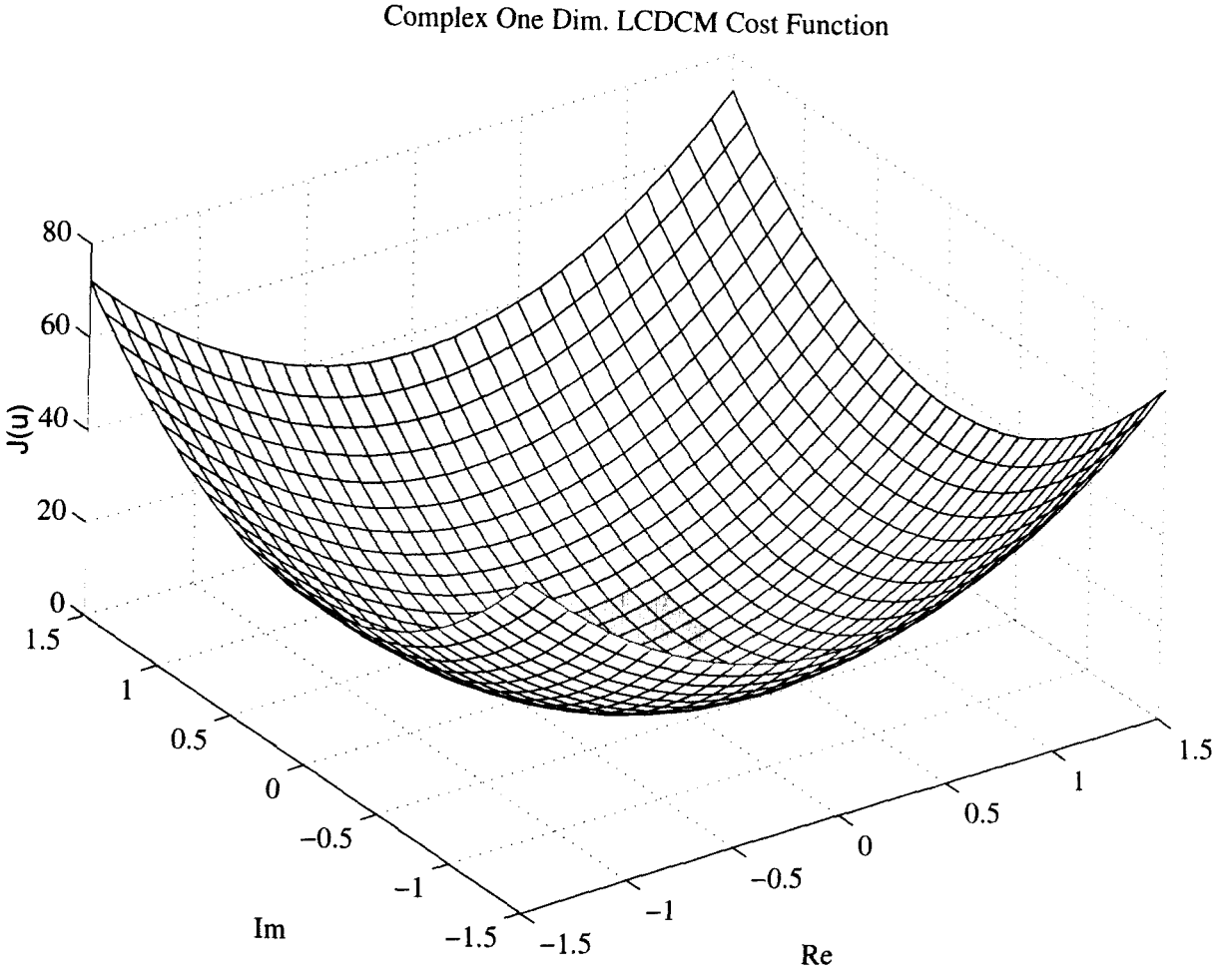

Figure 5.5: One Dimensional Complex LCDCM cost function surface.

$$J = \frac{1}{2}E\left[\left(\langle y, c_1^*\rangle^2 - \langle y_D, c_1^*\rangle^2\right)^2\right], \tag{5.74}$$

where c_1 is the multidimensional parameter which operates on y in the form of a linear transform.

We may now write the cost function in terms of the canonical representation of (5.43) and (5.44), ignoring the expected value[3]:

$$J = \frac{1}{2}\left(\langle y, s_1^* + x_1^*\rangle^2 - \langle y_D, s_1^* + x_1^*\rangle^2\right)^2. \tag{5.75}$$

Finding the multidimensional gradient of (5.75) we get

[3] As in the case of the LMS algorithm, we may do this. The reason for this is that in the execution of several iterations, the trajectory will be, on average, in the direction of the steepest descent.

$$\nabla J = 4\left(y\left\langle y, s_1^* + x_1^*\right\rangle - y_D\left\langle y_D, s_1^* + x_1^*\right\rangle\right)\left(\left\langle y, s_1^* + x_1^*\right\rangle^2 - \left\langle y_D, s_1^* + x_1^*\right\rangle^2\right) \tag{5.76}$$

Let us follow a similar approach whereby (5.49) is established. If we implement the linear constraint, the projection of the gradient in the direction orthogonal to s_1 is

$$\begin{aligned}\nabla J &= 4\left(\left[y - \left\langle y, s_1^*\right\rangle s_1\right]\left\langle y, s_1^* + x_1^*\right\rangle - \left[y_D - \left\langle y_D, s_1^*\right\rangle s_1\right]\left\langle y_D, s_1^* + x_1^*\right\rangle\right) \\ &\quad \left(\left\langle y, s_1^* + x_1^*\right\rangle^2 - \left\langle y_D, s_1^* + x_1^*\right\rangle^2\right)\end{aligned} \tag{5.77}$$

where the terms $\left[y - \left\langle y, s_1^*\right\rangle s_1\right]$ and $\left[y_D - \left\langle y_D, s_1^*\right\rangle s_1\right]$ are the projections or subspaces of y and y_D respectively for which the gradient stays orthogonal to s_1.

Denoting the matched filter responses for s_1 and $s_1 + x_1[i-1]$ as in (5.50) and (5.51), and the responses of the delayed signals with the subscript letter D, the adaptation rule of (5.77) is given by

$$x_1[i] = x_1[i-1] - \mu\left(\left[y[i] - Z_{MF}[i]s_1\right]Z[i] - \left[y_D[i] - Z_{MF_D}[i]s_1\right]Z[i]_D\right)\left(Z[i]^2 - Z[i]_D^2\right). \tag{5.78}$$

The finite dimensional vector implementation of the LCDCM algorithm is given by

$$\mathbf{x}_1[i] = \mathbf{x}_1[i-1] - \mu\left(\left[\mathbf{r}[i] - Z_{MF}[i]\mathbf{s}_1\right]Z[i] - \left[\mathbf{r}_D[i] - Z_{MF_D}[i]\mathbf{s}_1\right]Z[i]_D\right)\left(Z[i]^2 - Z[i]_D^2\right). \tag{5.79}$$

where the matched filter responses for $\mathbf{s}_1$ and $\mathbf{s}_1 + \mathbf{x}_1[i-1]$ are given by (5.54) and (5.55) and the responses of the delayed signals are again denoted by the subscript letter D.

5.4 PERFORMANCE OF THE LCCM AND LCDCM ALGORITHMS IN MULTIPATH FADING CHANNELS

The MMSE detector optimally combines multiple propagation paths, making it a very suitable receiver structure, given sufficient filter length to span all correlated paths. As we have seen, in single path environment, the LCCM and LCDCM detectors have the same vector weight solutions as the MMSE detector (assuming $\alpha > 1/4$ in the case of the LCCM detector). The single path vector weight solution of the MMSE, LCCM and LCDCM detector is given by (4.47) and [47] as

$$\bar{\mathbf{v}} = \mathbf{C}^{-1}\mathbf{s}_1 \tag{5.80}$$

The question now arises: How will the blind LCCM and LCDCM detectors fare in a multipath environment? Unfortunately, all multiple paths (except one) are suppressed as interference. The reason

for this is that in a multipath environment, $\mathbf{p}$ in (4.46) is no longer equal to $\mathbf{s}_1$. The vector $\mathbf{p}$ will now contain the contributions from the correlated parts of the delayed multipath components. Regardless of the multipath value of $\mathbf{p}$, the blind LCCM and LCDCM detectors will continue to extract only one path which correlates with $\mathbf{s}_1$.

Two ways have been proposed to allow the blind LCCM and LCDCM detectors to effectively combine the multiple paths:

1. The multipath channel can be estimated and used as the linear constraint for either the LCCM and LCDCM algorithms [60]. In this way all paths can be effectively combined.
2. A multi-channel LCCM (or LCDCM) algorithm as proposed by Mangalvedhe [47] can be used.

The former method is complex in that is requires singular value decomposition to estimate the multipath channel. The latter method uses several full detectors (channels) to extract each of the multiple paths. Adaptive weights are then used to optimally combine the outputs of the detectors. The fact that one needs a full detector to extract a single path also makes the multi-channel LCCM (or LCDCM) computationally expensive for a large number of paths. Although not discussed in this dissertation, it would be informative to compare the above mentioned two methods for multipath combination, both in terms of complexity, computational cost and performance.

Areas of possible further study could either be a search for a suitable cost function that will optimally combine multiple paths, or other methods to modify the LCCM or LCDCM algorithms which require less complexity than the above mentioned methods.

5.5 SUMMARY

This chapter contains much of the novel theoretical work attempted in this dissertation. The problem of blind multiuser detection utilizing the constant modulus algorithm is explored. An introductory section familiarizes the reader with the relatively recent history associated with blind multiuser detection, as well as all the research that has been attempted in this field.

The second section concerns itself with the thorough analysis of the LCCM cost function. For the first time, through rigorous analysis, a global condition for the convexity of LCCM cost function is derived. The nature of the stationary points are also examined. Subsequently, the LCCM algorithm is derived and presented.

The following section shows that the LCDCM criterion is the solution to non-convergence problems that, under certain circumstances, may plague the LCCMA. This is done by proving that a global minimum exist on the LCDCM cost function. Following the analysis of the LCDCM criterion, the LCDCM algorithm is derived.

In the final section, a qualitative analysis of the LCCM and LCDCM detector performance in a fading multipath channel is conducted. Methods to remedy shortcomings of these detectors in a multipath environment, are proposed.

CHAPTER SIX

SIMULATION RESULTS OF THE CM DETECTOR

In this chapter, implementation of a simulation platform and the results generated by this platform are discussed. In general, four linear detectors are compared, which include the single user matched filter detector, the MMSE detector, the LCCM detector and the LCDCM detector. The chapter is structured in the following way. The first section discusses the simulation platform setup for three different transmission channels which include the AWGN channel, the static multipath channel and the single- and multipath fading channels. Assumptions for each of these channels are also stipulated in this first section. In the following three sections, the simulation results for the three different types transmission channels are presented and discussed. Conclusions are summarized in the next chapter, along with proposals for further study in the relevant blind multiuser detection fields.

6.1 SIMULATION SETUP

The simulation platform was implemented in an object oriented C++ environment. The top level block diagram of the simulation setup is depicted in Figure 6.1. The transmitter was simulated by spreading random data using a seven length Gold sequence. Up to six users were simulated in this way. The channel that was simulated was either an AWGN channel, a static multipath channel, or a Rayleigh fading single- or multipath channel. Full details on the simulation of the mobile fading channel is given in Appendix B. The receiver was made to simulate either the matched filter, MMSE, LCCM and LCDCM detectors for comparative evaluation. The adaptive receivers all employed the steepest gradient descent algorithm on each of their respective cost functions.

Some general simulation assumptions need mentioning:

- A synchronous channel is assumed, i.e. all users are perfectly lined up.

- A baseband model is implemented in the simulation and perfect symbol (and chip) synchronization is assumed at the receiver.
- The system is sampled at one sample per chip. The signature waveform thus has a constant modulus taken sample by sample. The chip waveform is thus *not filtered*. The effect of interpolation and constant modulus filtering (such as root of unity (RU) filtering) is trivial, as the sequence can be seen as a "longer" constant modulus signature waveform. The processing gain introduced by this apparent increase in sequence length is negated by the correlation introduced between consecutive chips.
- Unless otherwise stated, all users have equal energy, i.e. $A_1, A_2, \ldots, A_K = 1.0$.
- Differential encoding is employed to negate the effect of phase ambiguity created by the blind detectors.
- The step size of each of the stochastic gradient algorithms is chosen such that all the adaptive detectors converge at approximately the same rate. As a result, $\mu_{MMSE} = 0.001$, $\mu_{LCCMA} = 0.00003$ and $\mu_{LCDCMA} = 0.00003$.

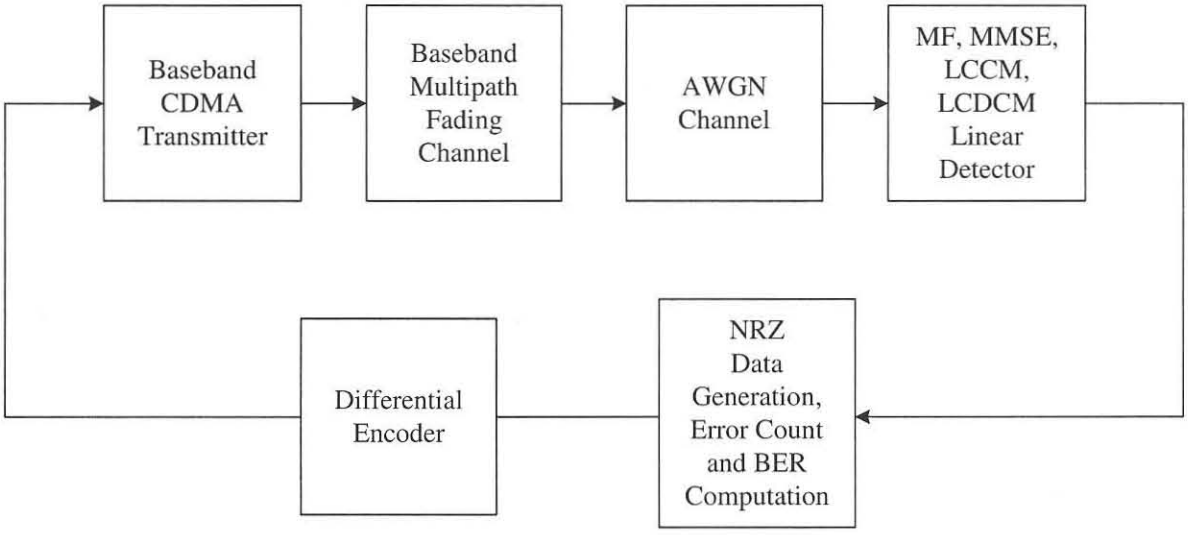

Figure 6.1: System block diagram of the simulation setup.

As we have seen in Chapter 3, the asynchronous case with K users is equivalent to the synchronous case with $2K - 2$ users. The first assumption is thus justified. To minimize simulation complexity and time duration, perfect carrier and clock synchronization is assumed.

A description of the channel types and assumptions associated with each type of channel are given in the following subsections.

6.1.1 SINGLE PATH AWGN CHANNEL

The AWGN channel was simulated by using the Marsaglia-Bray method to generate uniformly distributed samples, and subsequently using the Wichmann-Hill transformation to map the uniform distribution into a Gaussian distribution. These noise samples were then weighted with the noise standard deviation and added to the transmitted signal.

Assumptions - Single Path AWGN Channel

- The desired user sequence is assumed to be known at the receiver.
- The channel is assumed to have no channel distortion. (i.e. to be frequency flat)

6.1.2 STATIC MULTIPATH CHANNEL

The static multipath channel was simulated by a linear filter with a sampled impulse response equal to 0.86 at zero delay, 0.43 at one chip delay, and 0.26 at two chips delay. This channel is frequency selective with minimum phase. This means that the channel inverse can easily be approximated by a linear filter with finite length.

Assumptions - Static Multipath Channel

- The optimum linear inverse channel estimation in a mean square error sense is known at the receiver. This inverse channel estimation is used as a linear constraint in the case of the LCCM and LCDCM detectors.
- The channel is assumed to have minimum phase channel distortion. (i.e. to be frequency selective)

6.1.3 SINGLE- AND MULTIPATH FADING CHANNELS

The single- and multipath fading channels were simulated using Clarke's model as explained in Appendix B. The following assumptions were made regarding the simulation of the single- and multipath fading channels.

Assumptions - Single- and Multipath Fading Channel

- A three ray fading channel is assumed with a sampled multipath profile of 0.86 at zero delay, 0.43 at one chip delay, and 0.26 at two chips delay.
- All users experience the same fading channel, thereby simulating the CDMA downlink channel.
- A doppler frequency of 50Hz is assumed at a sampling frequency of 4M samples per second.
- The optimum linear instantaneous inverse channel estimation in the mean square error sense is known at the receiver. This inverse channel estimation is used as a linear constraint in the case of the LCCM and LCDCM detectors.
- The fading channel phase and amplitude is assumed to be known at the receiver.

Since the channel may have multiple paths that fade independently, the channel may readily assume a non-minimum phase form. In this case, the channel inverse may not be accurately approximated by a finite linear filter. This means that even after linear equalization at the receiver, much residual ISI may remain, thereby degrading the receiver performance. To approximate the channel inverse more accurately, a (non-linear) decision feedback structure will have to be considered.

6.2 PERFORMANCE IN AN AWGN CHANNEL

This section discusses the performance of the matched filter, MMSE, LCCM and LCDCM detectors in an AWGN channel. Even though such a channel is rarely encountered in a mobile environment, much insight may be gained on the operation of these detectors. As we have seen in the previous chapter, the optimum tap weight vector of the MMSE, LDCCM and LCDCM detectors assume the same values in the AWGN case. This notion is strongly enforced by the simulation results. The only information that is needed by the LCCM and LCDCM detectors, is the signature waveform of the desired user. In the case of the MMSE detector, a training sequence is needed.

The performance will largely be evaluated using two different, though related criteria. The first is signal to noise and interference ratio (SIR), and is the measure of how well the detector is able to cancel out interfering users. The second is bit error rate (BER), which evaluates the detector in terms of the number of bit errors made after reception. This is the most important performance measure, as it is the reliability of the transmission which most concerns digital communication engineers.

Figure 6.2 shows the signal to noise interference ratios of the LMS, LCCMA and LCDCMA versus time in an AWGN channel with a bit energy to noise spectral density ratio of $E_b/N_0 = 10$ dB. There are 6 simultaneous equal energy users, each employing a Gold sequence with a spreading factor of 7. From this figure it can be seen that the different adaptive linear multiuser detection techniques

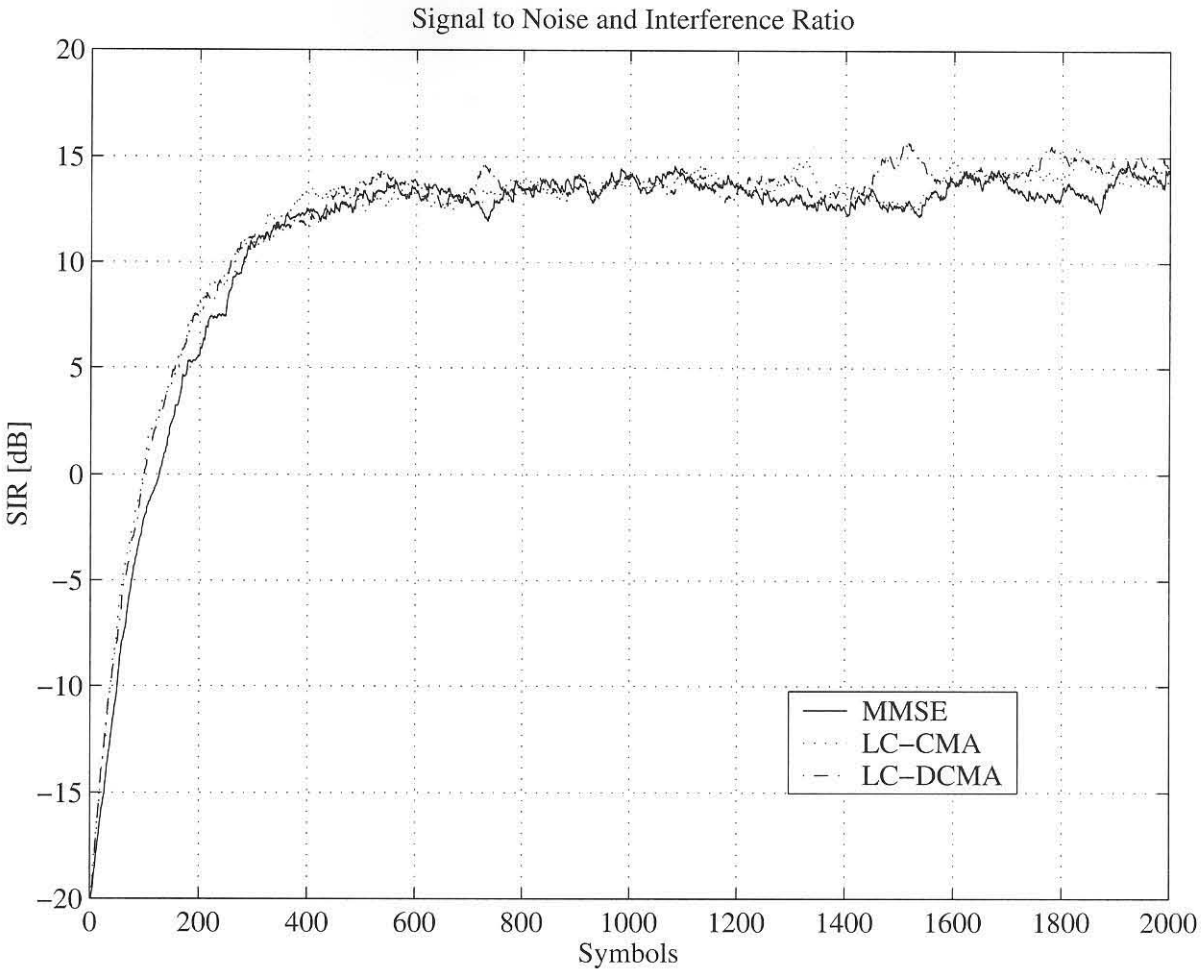

Figure 6.2: Signal to noise and interference ratios versus time of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the MMSE, LCCMA and LCDCMA detection techniques.

approach similar performance with similar speed of convergence. This intuitively satisfies the fact that the MMSE, LCCM and LCDMA have the same tap vector weight solutions. Note that in this case $A_1^* A_1 \geq \alpha/4$, which insures convergence of the LCCMA.

In Figure 6.3, we have that $A_1^* A_1 < \alpha/4$, which means that the LCCMA will not converge to the desired minimum. This is evident in the figure, in that the SIR in the case of the LCCMA decreases as time passes. Note that this condition may readily be encountered in a automatic gain controlled (AGC) uplink channel where all the users fade independently. If the fading on the desired user is severe, while it is not on the other users, the desired user power level may be below the threshold $\alpha/4$. In the downlink channel on the other hand, all users have the same amplitude, and the AGC will keep all the amplitudes above the threshold. From Figure 6.3, we can see that the LCDCM algorithm

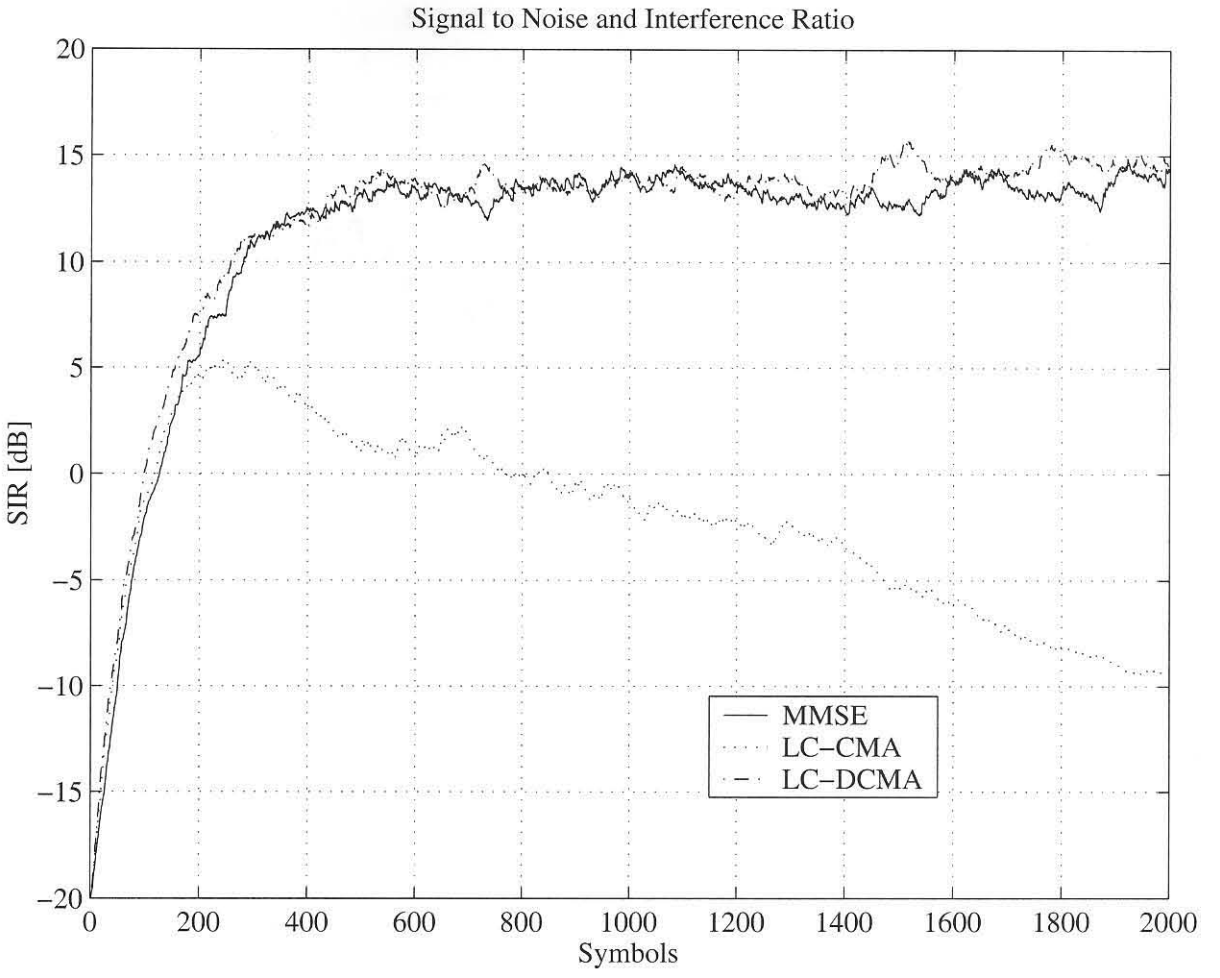

Figure 6.3: Signal to noise and interference ratios versus time of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the MMSE, LCCMA and LCDCMA detection techniques. In this case $A_1^* A_1 < \alpha/4$.

is resistant to this condition. This observation reinforces the fact as derived in Chapter 5, that the LCDMA has a global minimum regardless of the value of the desired user amplitude.

Figure 6.4 shows the response of the relevant detectors to a strong user (34dB) powering on in the channel. Note that the MMSE detector is more severely affected by this event than the LCCM and LCDCM detectors. This is easily explained by the fact the the matched filter component of the LCCM and LCDCM detectors is not affected by the channel disruption. Only the adaptive part of these detectors are affected. Consequently, since the MMSE detector is purely adaptive, it is more sensitive to changes in the channel. It must be mentioned that in the multipath case where the channel is adaptively estimated and employed in the linear constraint, for the same reason as the MMSE

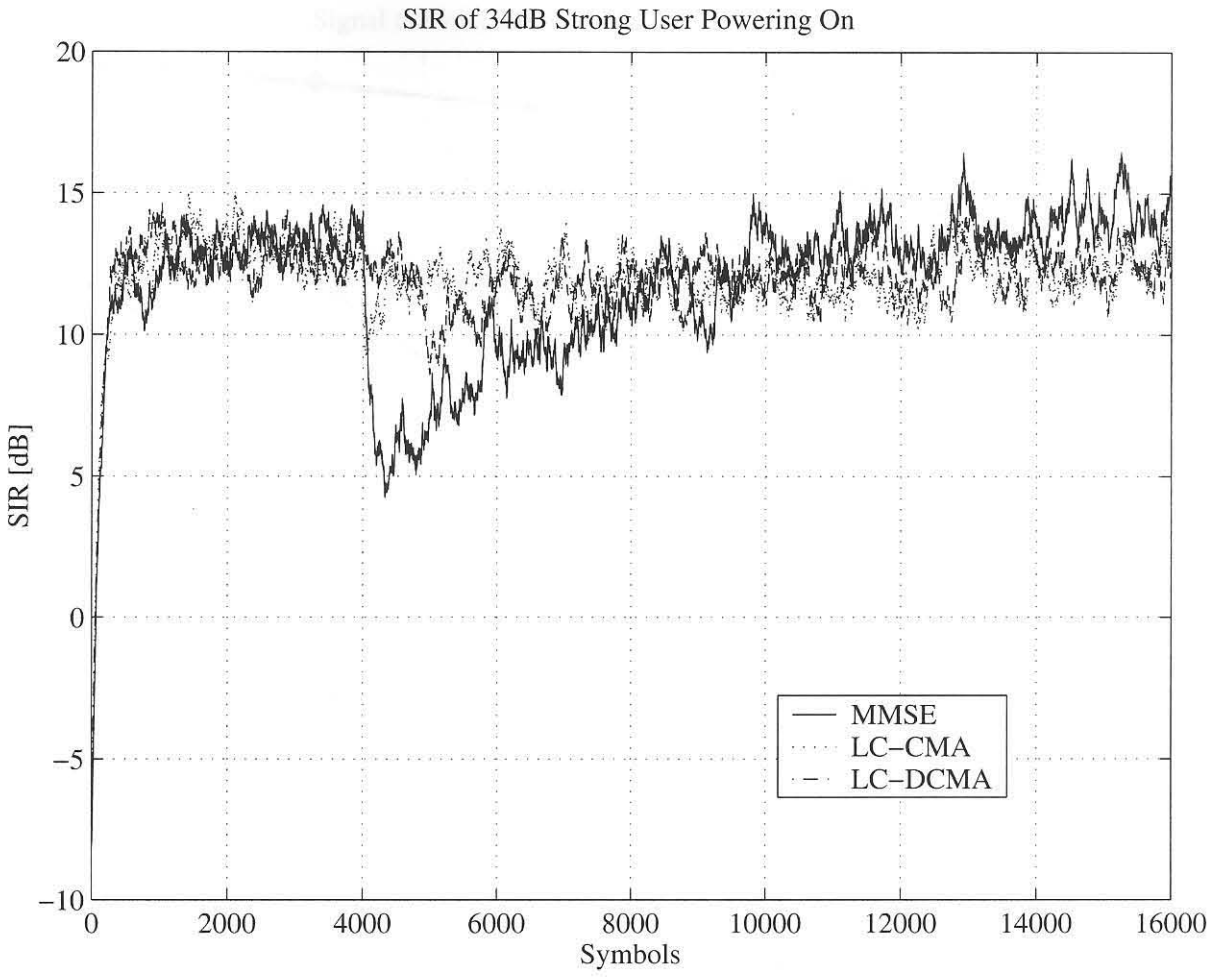

Figure 6.4: Signal to noise and interference ratios versus time of a CDMA system with 3 users, and a 34dB strong fourth user powering on at time $t = 4000$ symbols.

detector, the LCCM and LCDCM detectors' performance will also deteriorate if a strong user powers on.

Another important consideration, is how these multiuser detection techniques perform as the number of users increase. Figure 6.5 shows the SIR versus the number of interfering users. Here the matched filter, LMS algorithm, LCCMA and LCDCMA are compared in a AWGN environment with a bit energy to noise spectral density ratio of $E_b/N_0 = 10$ dB. It is evident that the matched filter ignores the contribution of the interferers, and fares rather poorly as the number of users increase. All the adaptive detector are able to cancel out the interference, only at a slight penalty in SIR as the number of users increase. Also here, the MMSE, LCCM, and LCDCM criteria exhibit comparable performance due to the fact that they have the same tap weight vector solutions.

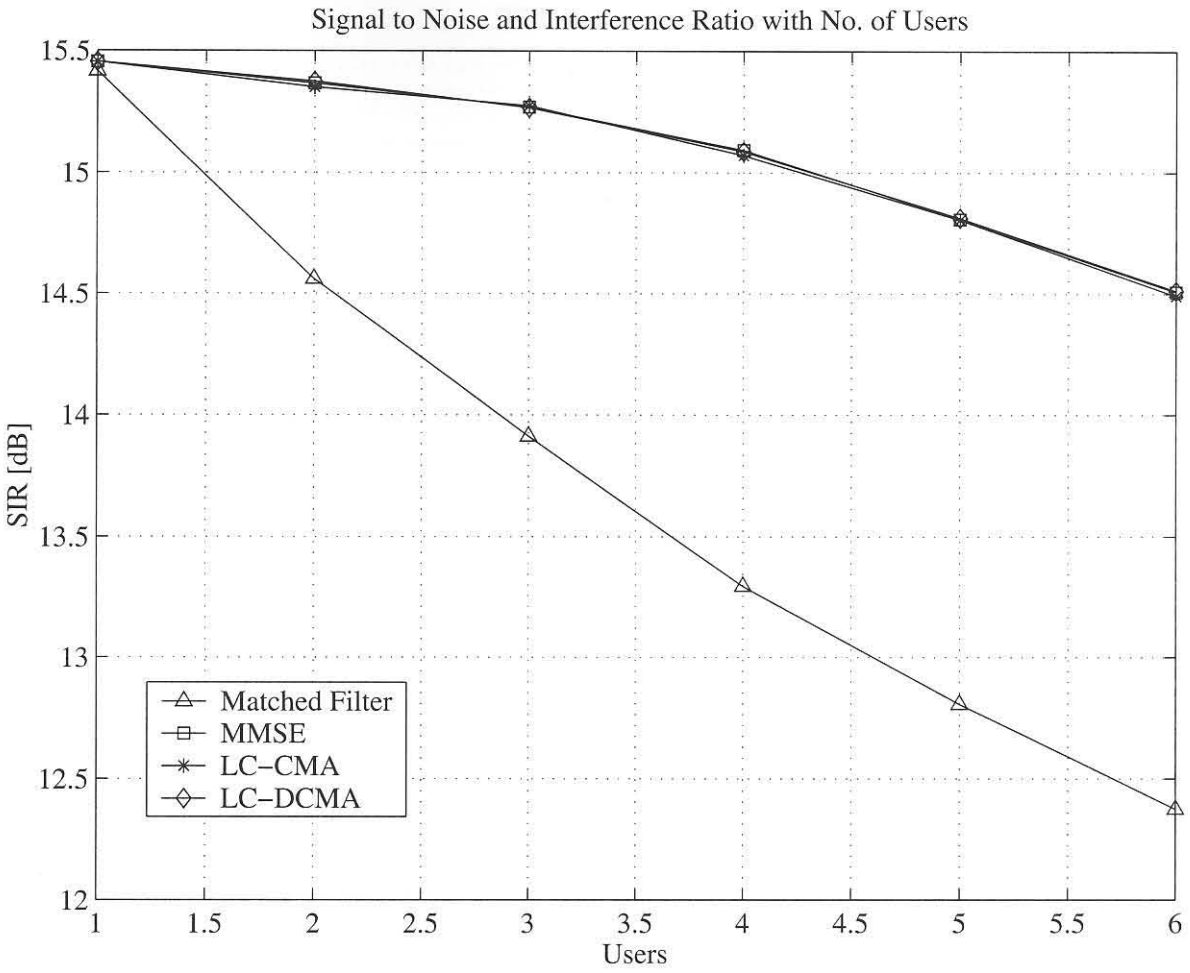

Figure 6.5: Signal to noise and interference ratios versus no. of users of a CDMA system with a spreading factor of 7 and $E_b/N_0 = 10$ dB in an AWGN channel.

Figure 6.6 shows that due to the increased number of uncancelled interferers, the matched filter detector exhibits a significant increase in BER as the number of users increase. Concerning the adaptive detectors, due to only a slight drop in SIR as the number of interferers grow in Figure 6.5, we can expect a slight increase in BER in Figure 6.6. This is readily verified by comparing the two figures. This once again supports the notion that there exists a strong relation between SIR and BER.

Since we are using the desired user's signature waveform as the linear constraint for the LCCM and LCDCM detectors, it might be informative to see how inaccuracies in the constraint affect the performance of these detectors. This will be of interest in multipath channels, as inaccuracies in channel estimation adversely affects the operation of the LCCM and LCDCM detector. Note that it is well documented [20] [60], that the LCCM detector is much more robust to these inaccuracies that the

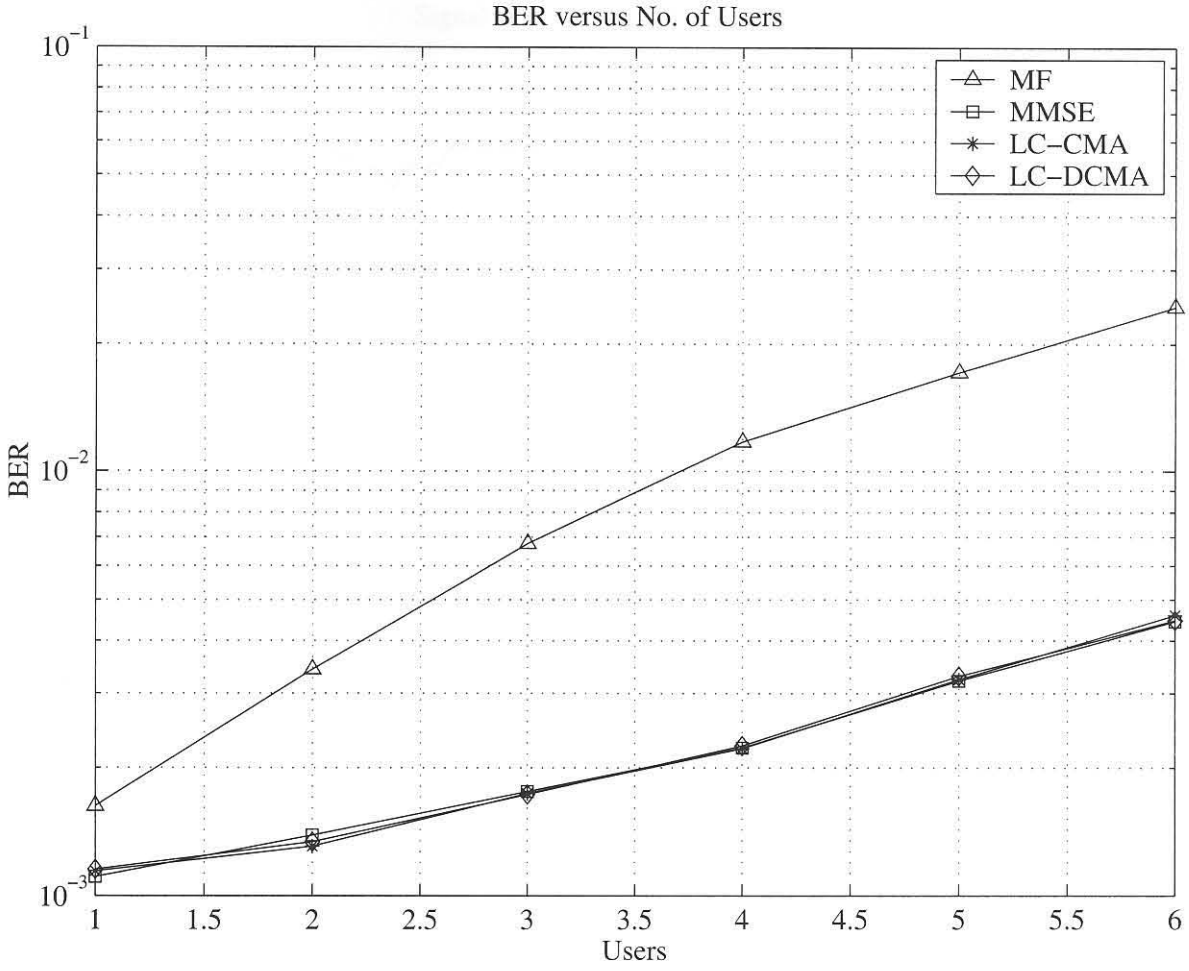

Figure 6.6: Bit error rate versus no. of users of a CDMA system with a spreading factor of 7 and $E_b/N_0 = 10$ dB in an AWGN channel.

LCMV detector of Honig [3]. In Figure 6.7 it can be seen that the LCCM and the LCDCM detectors appear to achieve the same robustness to signature waveform mismatch with a variance of 0.1, though some loss in SIR is inevitable. This may be due to two factors: the fact that the receiver is not a perfect matched filter with respect to the transmitter and that the desired component is not sufficiently protected from being cancelled out due to an imperfect constraint.

The final figure in this section (Figure 6.8) depicts the simulation BER of an AWGN channel with 6 users and Gold sequences of length 7. The high channel load clearly renders the matched filter ineffective. The MMSE, LCCM and LCDCM detectors are effective at mitigating the multiuser interference problem to a large extent, but still does not approach the single user bound. The only detector that is able to approach the single user bound is the ML sequence detector, but is disqualified due to

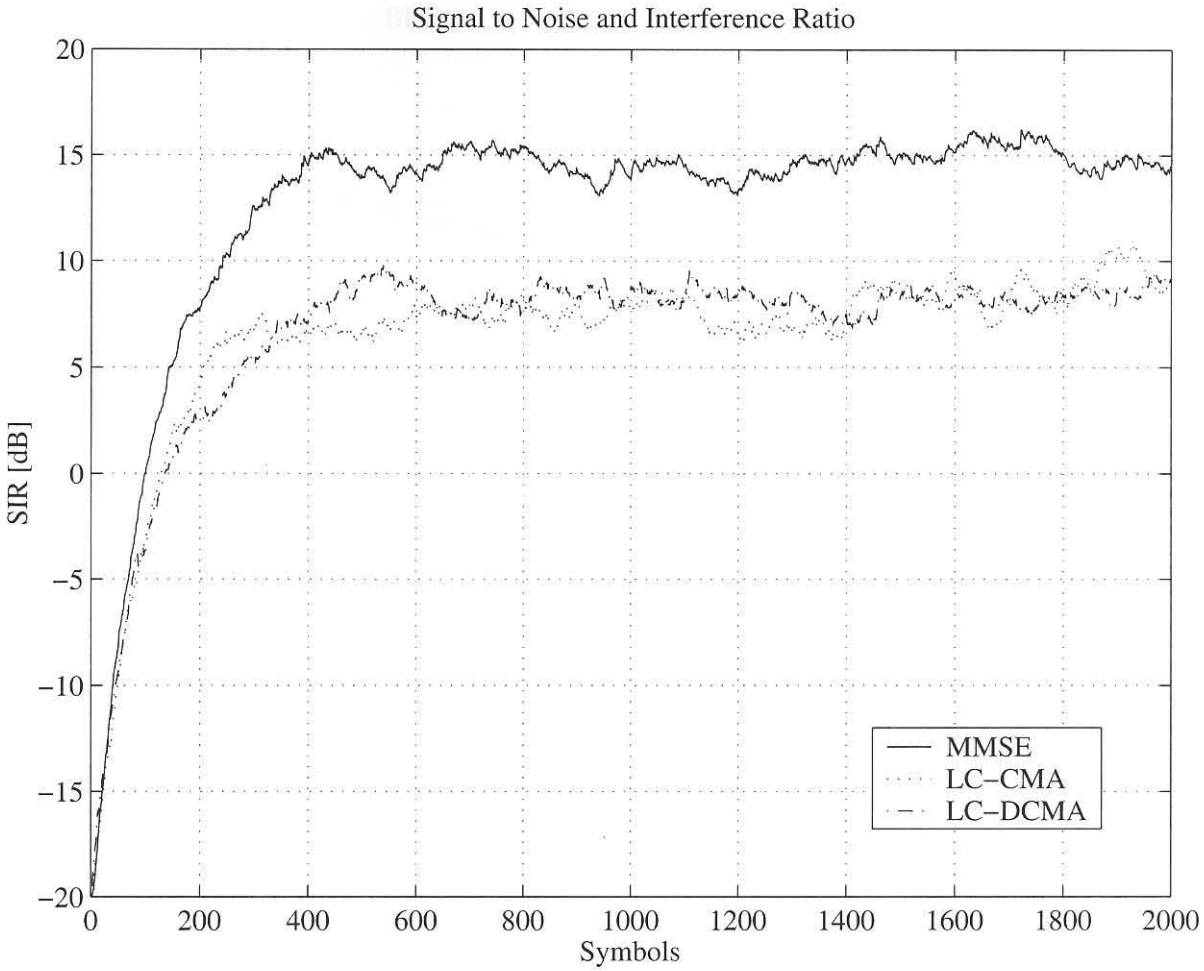

Figure 6.7: Signal to noise and interference ratios versus time of a CDMA system with K = 6 users and a SF = 7 in an AWGN channel. The plot shows the performance for a code mismatch with mismatch variance of 0.1.

its complexity, which exponentially increases with the number of users.

6.3 PERFORMANCE IN A STATIC MULTIPATH CHANNEL

We will now consider the performance of the matched filter, MMSE, LCCM and LCDCM detector in a static multipath channel. Here we choose a minimum phase channel, of which the inverse can be accurately approximated by a finite length linear filter. This means that the inverse channel impulse response rapidly decays to zero within the length of the receiver filter. The channel is chosen such that the linear receiver will be able to adaptively combine the multiple paths of the frequency selective channel. Channels with this type of impulse response are known as channels with *mild intersymbol*

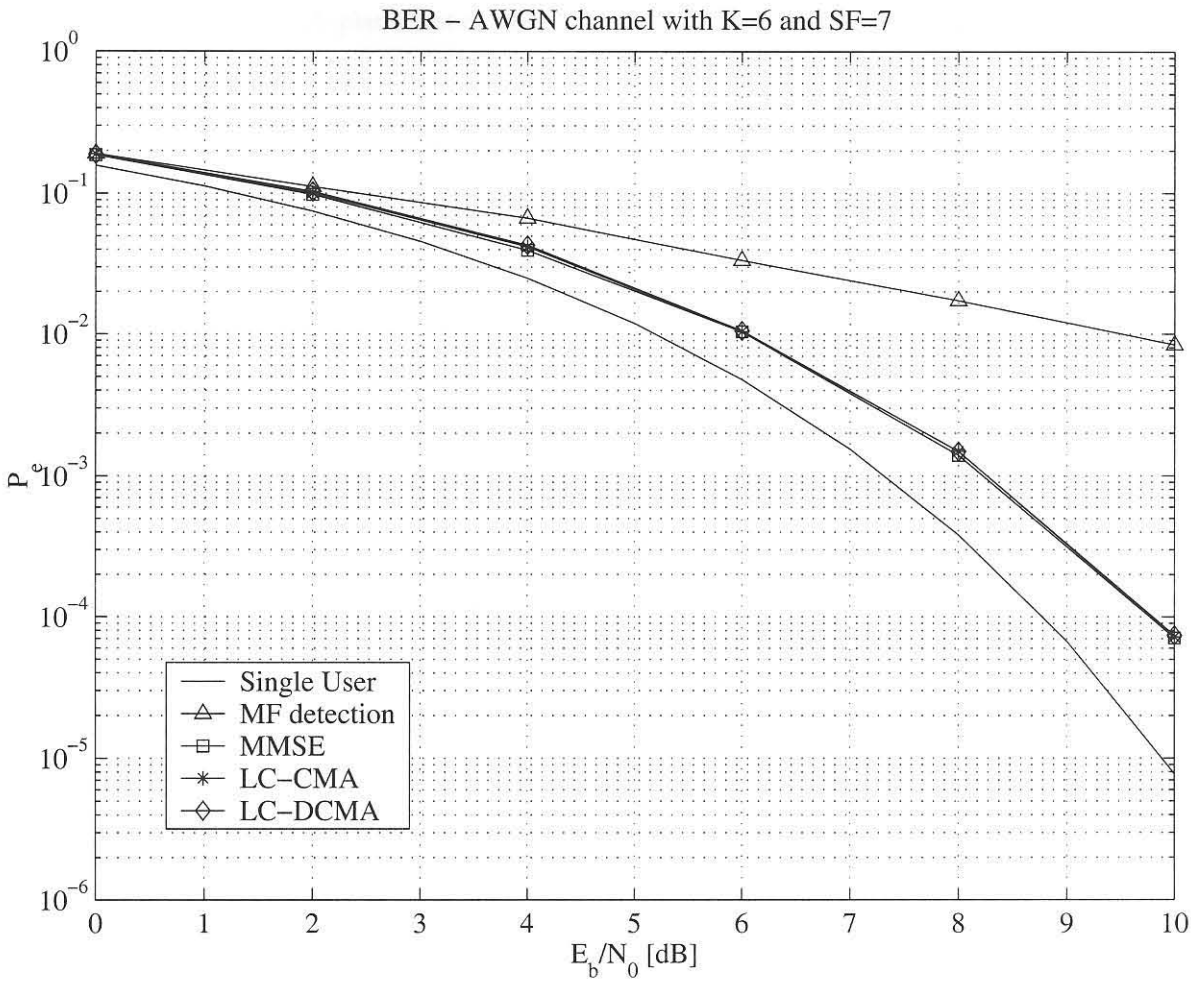

Figure 6.8: BER of a CDMA system with 6 users and a spreading factor of 7 in an AWGN channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques.

interference. We have seen in Chapter 5 that the LCCM and LCDCM detectors will attempt to suppress the multiple paths, instead of combining them. For these detectors to be able to combine the paths, accurate channel estimation is needed to find the optimum channel plus noise inverse, and use it (convolved with the desired signal waveform) as the linear adaptation constraint. In this way the effect of the multipath channel is negated, and the multiple paths are optimally combined. The multiuser interference is then cancelled out on the subspace orthogonal to this modified constraint. Once again it must be stated that it is only possible if the channel (plus noise) inverse can be accurately approximated by a finite linear filter. The MMSE detector on the other hand, will automatically find the best linear inverse channel plus noise in the mean square error sense that is also able to cancel out the multiuser interference.

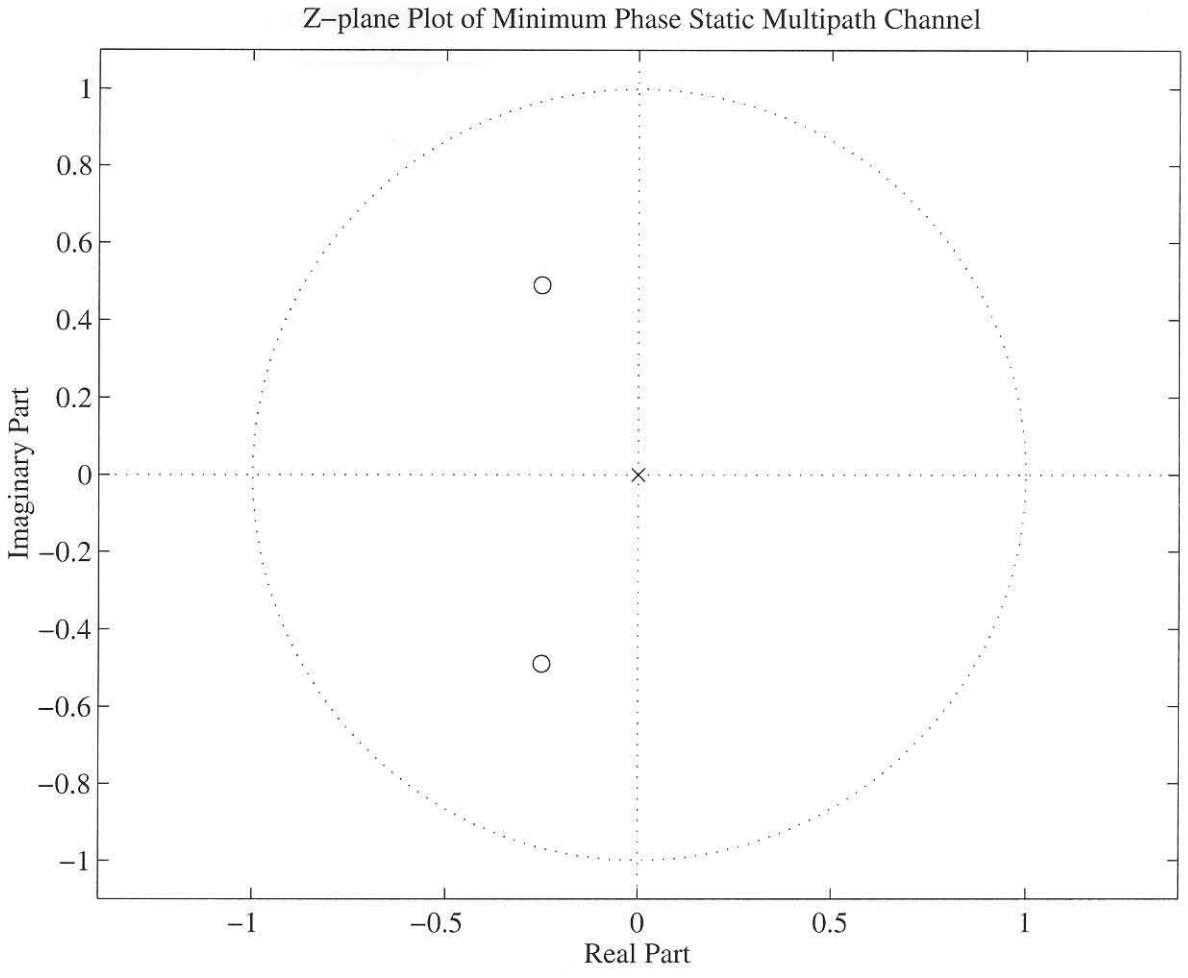

Figure 6.9: Z-plane plot of a minimum phase static 3-ray multipath channel. The multipath profile consists of 0.86 at zero delay, 0.43 at 1 chip delay and 0.26 at 2 chips delay.

In the simulation, perfect channel knowledge is assumed. Many authors propose methods to achieve accurate channel estimation ([60], [15], [68]), but it is beyond the scope of this dissertation.

The z-plane plot of the simulated channel with impulse response $\{0.86, 0.43, 0.26\}$ is given in Figure 6.9. The channel only has two singularities (zeros) within the unit circle, clearly indicating that it is minimum phase. The frequency response of the multipath channel is depicted in Figure 6.10, from which it can be readily seen that there exists no zero in the spectrum. The inverse of the channel can thus easily be obtained without needing a pole with infinite gain. Once again this supports the fact that a linear (feed forward) filter is sufficient to negate the effect of the channel, by forming the inverse frequency response of the channel.

The simulated BER is depicted in Figure 6.11. From this figure it can be seen that the LCCM and

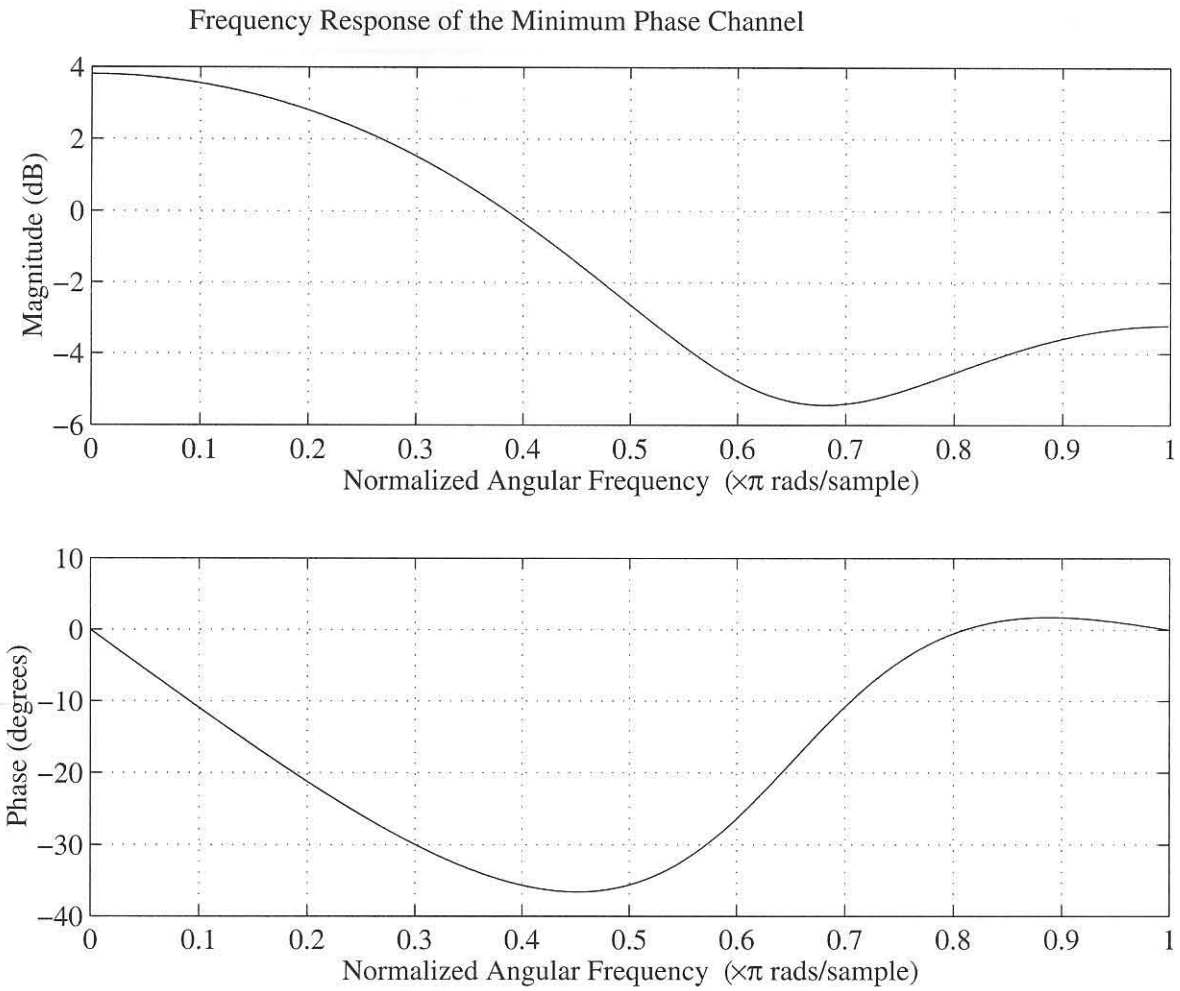

Figure 6.10: Frequency response of the static three-ray multipath channel of which the z-plane representation is shown in Figure 6.9.

LCDCM detectors with optimum linear channel estimation are just as effective as the MMSE detector in simultaneously combating multipath and multiuser interference. As can be expected, the matched filter has no chance in effectively demodulating a signal that has multipath in addition to being saturated with users.

6.4 PERFORMANCE IN RAYLEIGH FADING SINGLE- AND MULTIPATH CHANNELS

Before we consider the multipath fading case, we will first consider the single path fading channel. In this case, the fading is frequency flat, i.e. no channel distortion is introduced. As explained in

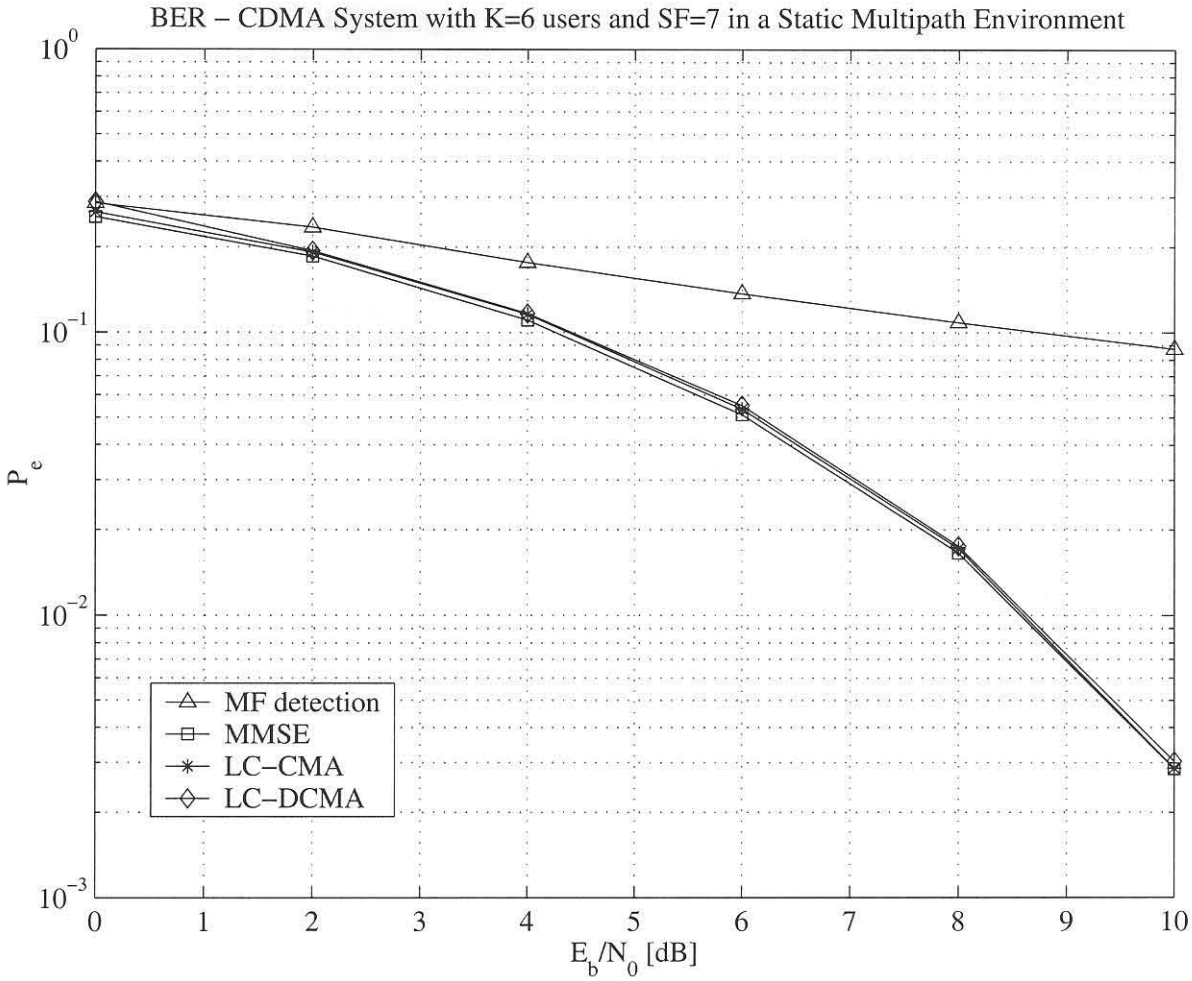

Figure 6.11: BER of a CDMA system with 6 users and a spreading factor of 7 in a static 3-ray multipath channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques. The multipath profile consists of 0.86 at zero delay, 0.43 at 1 chip delay and 0.26 at 2 chips delay.

Appendix B, the amplitude envelope of the fading channel assumes a Rayleigh distribution. The doppler frequency of the fading channel is chosen to be 50Hz in a system sampled at 4M samples per second. At this sampling rate, the fading is rather slow, and can easily be followed by an adaptive receiver employing a stochastic gradient descent algorithm. We assume that all users experience the same amount of fading, thereby simulating the CDMA downlink channel. The deep fades introduced by such a channel introduces a large burst of errors, severely degrading the effective SNR and consequently the average BER. It is for this reason that Figure 6.12 shows a modest improvement in BER compared to the matched filter receiver. The periods of deep fades tend to make the AWGN dominant over the multiuser interference, thereby reducing the margin of performance increase that can be achieved by cancelling out the multiuser interference.

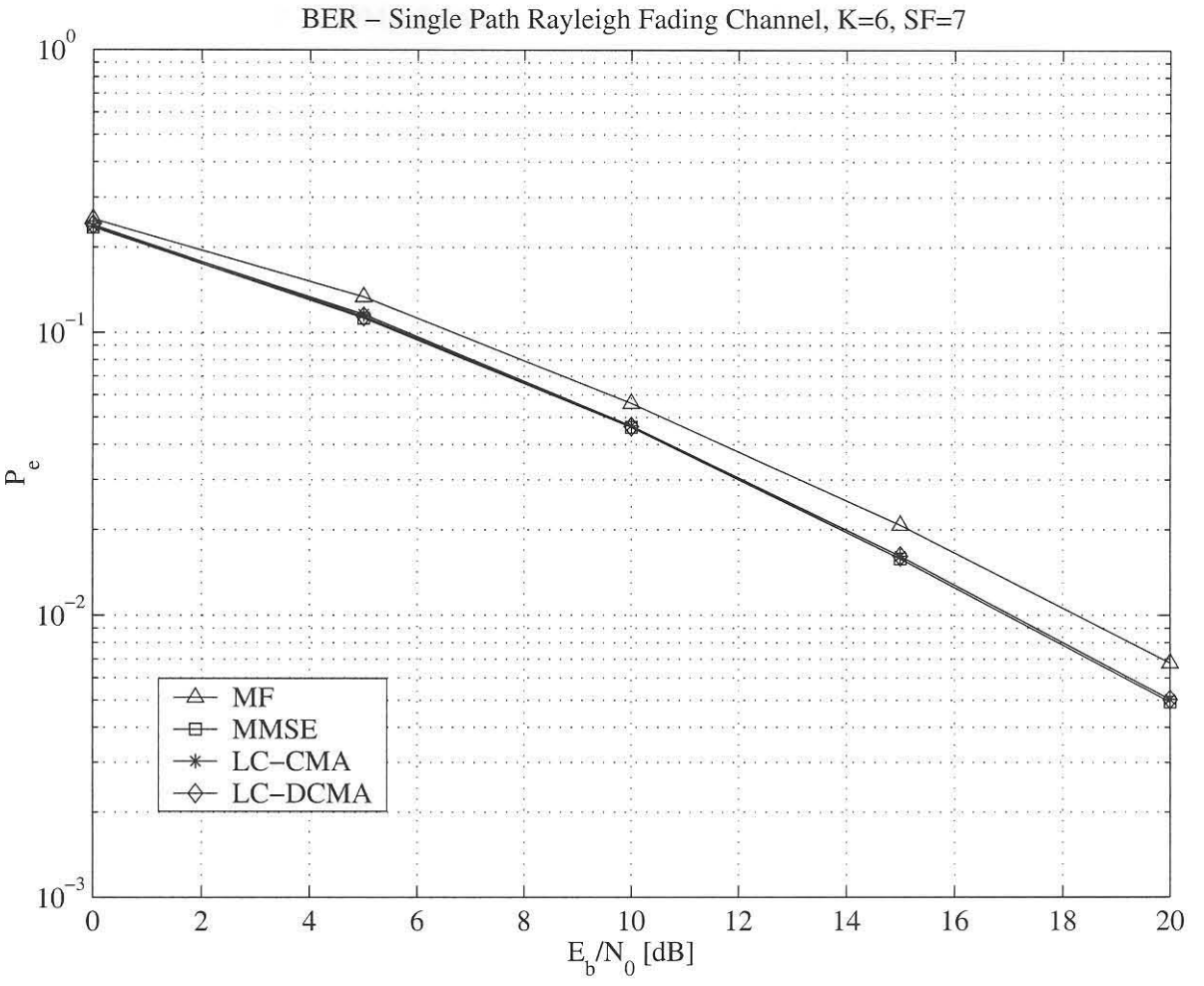

Figure 6.12: BER of a CDMA system with 6 users and a spreading factor of 7 in a Rayleigh fading single-path channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques.

As mentioned in the previous section, it is possible to effectively combine multiple paths of channels with *mild* ISI by using a linear filter. In a mobile multipath fading channel, the amount of ISI may vary between mild and severe. In the case of severe ISI (a non minimum phase channel), *no* finite length linear filter is able to effectively combine the multiple paths and negate the effect of ISI. In this case we will need to employ a non-linear or *decision feedback* structure to estimate the channel inverse (zero forcing criterion) or to estimate the channel plus noise inverse (MMSE criterion). This implementation, however, is beyond the scope of this dissertation, but warrants some further investigation.

The effect of a multipath fading channel on the relevant linear multiuser structures with linear channel

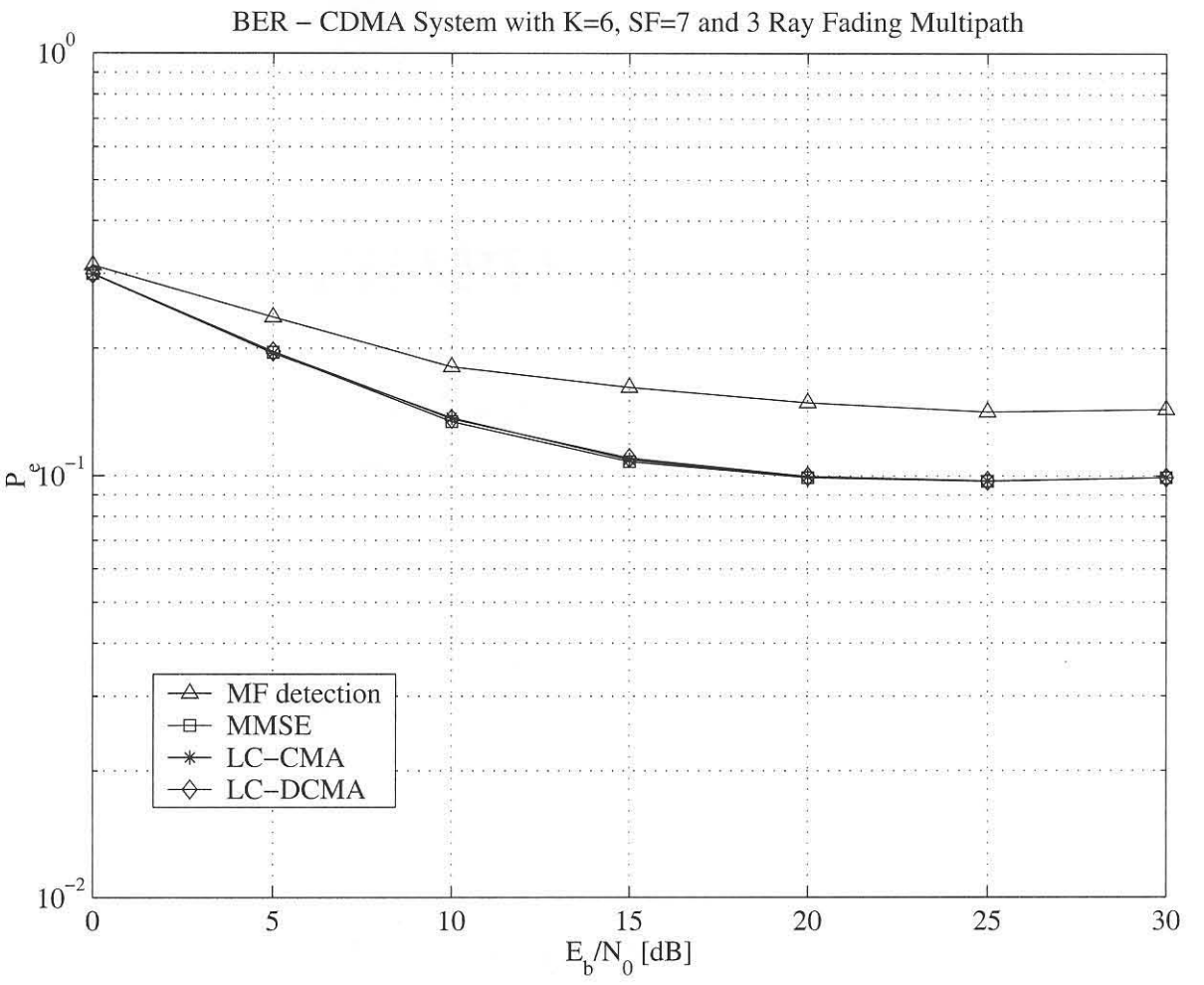

Figure 6.13: BER of a CDMA system with 6 users and a spreading factor of 7 in a 3-ray Rayleigh fading multipath channel using the matched filter, MMSE, LCCMA and LCDCMA detection techniques.

estimation is depicted in Figure 6.13. Due to the fact that a finite linear filter is unable to accurately model the mobile channel inverse, some residual ISI remains uncancelled. Looking at the BER curve beyond 25dB, the BER seems to increase. This strongly resembles the closed eye condition as depicted in Figure 3.13. The ISI causes the signal to move to the wrong side of the decision boundary. In this case, the addition of noise benefits the BER, as it may move the signal back to the correct side of the decision threshold.

CHAPTER SEVEN

FINAL SUMMARY AND CONCLUSIONS

Chapter seven concludes this dissertation, and contains the summaries and conclusions derived from all the previous chapters. The novel work attempted in this dissertation is clarified and its significance is discussed here. The first section describes significant discoveries and concepts with respect to the theoretical part of this dissertation. The second section summarizes practical simulation results and evaluations, and discusses the implications of these results. The final section proposes topics which may be considered for future work and research.

7.1 THEORETICAL SUMMARY AND CONCLUSIONS

In this dissertation the constant modulus type (LCCM and LCDCM) detectors are analyzed within the context of linear detectors. Linear detectors perform a linear vector operation on the received signal vector. The adaptive linear detector solves many of the complexity and assumed knowledge (Figure 7.1) issues associated with many of the other multiuser detector structures. Chapter four starts by characterizing the optimum linear multiuser detector in terms of multiuser efficiency. The optimum linear multiuser detector suffers from the same penalty as the decorrelating (or zero forcing) detector in the low SNR region, as it ignores the contribution of the noise. The MMSE and related blind (LCMV, LCCM, LCDCM) detectors allows some residual multiuser interference to remain in order to attain optimal performance with respect to AWGN and multiuser interference. In this dissertation, the real valued model in Verdu [31] is extended to encompass complex values. Differentiation with respect to a complex vector [57] is extended to differentiation with respect to a complex matrix in Appendix C. The complex gradient of the MMSE detector is consequently derived, and the MMSE detector is analyzed in the complex domain. Uniquely, in this dissertation, the model in [31] is also extended to model a multipath channel in a similar manner as in [47].

	Single-user Matched Filter	Decorrelator	MMSE	Adaptive MMSE	Blind MMSE (LCCMA, LCDCMA)	Decision Driven (SIC,PIC,etc)
Code of desired user	●	●	●		●	●
Timing of desired user	●	●	●	●	●	●
Received amplitudes			●			●
Noise level			●			
Code of interfering users		●	●			●
Timing of interfering users		●	●			●
Training seq. of desired user's data				●		

Figure 7.1: Knowledge needed for the different types of multiuser detection schemes.

Much novel mathematical analysis of the LCCM and LCDCM detectors is achieved in Chapter five. For the first time, a global condition is derived for the convexity of LCCM cost function. Some previous authors evaluated the LCCM cost function at the desired stationary point or vector solution, without considering any other possible solutions. This is an error, as any of the other stationary points may or may not be an undesired global minimum. On the other hand, the LCDCM cost function is shown to have a single stationary point, which is also a minimum. The convexity of the LCDCM cost function needs not even be considered, since the only stationary point is the global minimum. Using the same process as in Verdu [31], the stochastic gradient algorithms for the LCCM and LCDCM criteria are also derived. At the end of Chapter five, the effect of multipath on the LCCM and LCDCM detectors is qualitatively discussed. Under normal operation, the MMSE detector can combine multipaths, while the LCCM and LCDCM detectors merely attempt to cancel it out. If inverse channel plus noise estimation can be used within the linear constraint, the LCCM and LCDCM detectors can effectively combine the multiple paths, providing that the multipath channel plus noise inverse can be accurately modelled within the length of the detector. The multipath combination is achieved by convolving the estimated channel inverse plus noise with the desired user signature waveform. This is then used as the modified linear constraint in the LCCM and LCDCM detectors.

7.2 SIMULATION SUMMARY AND CONCLUSIONS

One can conclude from the simulation results that the basic LCCM and LCDCM detectors are effective in mitigating multiuser interference in AWGN and frequency flat fading channels. In all the simulation results, it is apparent that the performance of the LCCM[1] and LCDCM detectors approximate that of the MMSE detector. Even as the number of users increase, the LCCM and LCDCM detectors' performance match that of the MMSE detector. The SIR measurements in the case of $A_1^* A_1 < \alpha/4$ shows that while the LCCM detector fails, the LCDCM detector performs on par with the MMSE detector. This is consistent with the theoretical derivations done in Chapter five, where it is shown that the LCCM Hessian matrix becomes negative definite at the desired stationary point, indicating a local maximum. On the other hand, the LCDCM cost function has a global minimum, irrespective of desired user amplitude. Concerning BER performance, the LCCM, LCDCM and MMSE detectors show a massive improvement over the matched filter in the AWGN channel. There is still some small improvement to be gained when compared with the single user bound. This, however, is only about 1dB at an E_b/N_0 of 10dB. We can thus conclude that for the complexity of the adaptive MMSE type linear detector, the performance that is gained when compared with the optimum (non-linear) detector is excellent. The optimum non-linear detector for the same 6 user channel will require a trellis of 64 states, which is a complex detector for only 6 users. In a CDMA channel with 20 users the number of trellis states for the optimum detector will increase to over a million states, which is simply impossible to implement. If a linear detector were to be employed for the same channel, the complexity will only triple compared to the six user linear detector case.

In non fading frequency selective channels that are well behaved (mild ISI and minimum phase), the linear structure of the LCCM and LCDCM detectors exploits the multiple paths effectively, even if an accurate linear channel plus noise estimate is employed in the linear constraint. In this case, the performance of the LCCM and LCDCM detector approaches that of the best linear detector, viz. the MMSE detector. In multipath fading channels, all the linear detectors[2] are effective at cancelling out multiuser interference, but are insufficient at optimally combining the multiple paths. This is because the multipath mobile fading channel plus noise may assume a form that cannot be accurately inversed by using a finite linear structure. In this case, the use of a non-linear decision feedback structure may warrant further investigation.

[1] Here it is assumed that $A_1^* A_1 \geq \alpha/4$.

[2] In this case the LCCM and LCDCM detectors are assumed to have for linear constraints, the best linear inverse channel plus noise estimation in a mean square error sense.

7.3 PROPOSALS FOR FURTHER RESEARCH

From the multipath fading channel results obtained from this dissertation, the question that is raised is if it's possible to do joint blind multiuser detection and fading multipath equalization in a single receiver structure. Previous authors [15, 60, 68] proposed the use of channel estimation in order to obtain the multipath linear constraint. These methods are cumbersome, and it means that two adaptive structures are needed for joint multiuser detection and multipath combination. Furthermore the channels have to be well behaved (minimum phase) for the inverse to be approximated by a finite linear filter. A second method for joint multiuser detection and multipath combination is the multichannel detector proposed by Mangalvedhe [47]. This structure has high complexity, especially where many multipaths are concerned. In obtaining a single simpler joint fading multipath combiner and multiuser detector, it is obvious that some sort of non-linear decision feedback detector should be employed to be able to equalize non-minimum phase channels. It would be informative and advantageous to pursue such an avenue in future research. In this context, fast adaptation algorithms and fractionally spaced structures can also be examined to be able to cope with fast fading multipath channels [57].
Implementation concerns in DSP or FPGA, are of paramount importance when considering multiuser detectors. It should be informative to evaluate the behavior of fixed point implementations of blind adaptive multiuser detector structures, as well as any limitations revealed in this regard.
Application of the blind constant modulus algorithms to specific existing DS-CDMA systems employing constant envelope complex spreading sequences, warrants further investigation [5,69]. These systems offer good fading channel performance along with excellent non-linear amplification performance figures.

REFERENCES

[1] A.Klein, "Multi-user detection of CDMA signals - algorithms and their application to cellular radio," tech. rep., Fortschr.-Ber. VDI Reihe 10, Nr. 423. Dusseldorf: VDI-Verlag 1996, 1996.

[2] S. Verdu, "Minimum probability of error for asynchronous gaussian multiple access channels," *IEEE Trans. on Info. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[3] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. on Info. Theory*, vol. 41, pp. 944–960, July 1995.

[4] C. Johnson, P. Schniter, T. Endres, J. D. Behm, D. Brown, and R. Casas, "Blind equalization using the constant modulus algorithm," *Proc. of the IEEE*, Oct. 1998.

[5] M. Jamil, L. P. Linde, J. E. Cilliers, and D. J. van Wyk, "Comparison of complex spreading sequences based on filtering methods and mean square correlation properties," *Trans. of the SAIEE*, vol. 89, pp. 90–97, Sept. 1998.

[6] L. Staphorst, M. Jamil, and L. P. Linde, "Performance of a synchronous balanced QPSK CDMA system using complex spreading sequences in AWGN," in *Proc. of Africon '99*, pp. 215–220, Cape Technicon, Sept. 1999.

[7] G. Woodward and B. S. Vucetic, "Adaptive detection for DS-CDMA," *Proc. of the IEEE*, vol. 86, pp. 1413–1434, July 1998.

[8] J. Laster and J. Reed, "Interference rejection in digital wireless communications," *IEEE Signal Processing Magazine*, pp. 37–60, May 1997.

[9] A. Deul-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Personal Communications*, pp. 46–58, April 1995.

[10] Z. Xie, C. K. Rushforth, and R. T. Short, "Multiuser signal detection using sequential decoding," *IEEE Trans. Commun.*, vol. 38, pp. 578–583, May 1990.

[11] P. Alexander, L. Rasmussen, and C. Schlegel, "A linear receiver for coded multiuser CDMA," *IEEE Trans. Commun*, vol. 45, pp. 605–610, May 1997.

[12] L. Wei, L. Rasmussen, and R. Wyrwas, "Near optimum treesearch detection schemes for bit-synchronous multiuser CDMA systems over gaussian and two-path rayleigh fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 691–700, June 1997.

[13] X. Li and H. Fan, "Direct blind multiuser detection for CDMA in multipath without channel estimation," *IEEE Transactions on Signal Processing*, vol. 49, pp. 63–73, Jan. 2001.

[14] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.

[15] M. Torlak and G. Xu, "Blind multiuser channel estimation an asynchronous CDMA systems," *IEEE Transactions on Signal Processing*, vol. 45, pp. 137–147, Jan. 1997.

[16] X. Wang and H. V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.

[17] M. K. Tsatsanis, "Inverse filtering criteria for CDMA systems," *IEEE Transactions on Signal Processing*, vol. 45, pp. 102–112, Jan. 1997.

[18] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Transactions on Signal Processing*, vol. 46, pp. 3014–3022, Nov. 1998.

[19] J. K. Tugnait, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Goddard cost function," *IEEE Transactions on Signal Processing*, vol. 45, pp. 268–271, Jan. 1997.

[20] J. Miguez and L. Castedo, "A linearly constrained constant modulus approach to blind adaptive multiuser interference suppression," *IEEE Commun. Lett.*, vol. 2, pp. 217–219, Aug. 1998.

[21] C. Xu and G. Feng, "Comments on a linearly constrained constant modulus approach to blind adaptive multiuser interference suppression," *IEEE Commun. Lett.*, vol. 4, pp. 280–282, Sept. 2000.

[22] Z. Tang, Z. Yang, and Y. Yao, "Closed-form analysis of linearly constrained CMA based blind multiuser detector," *IEEE Commun. Lett.*, vol. 4, pp. 273–276, Sept. 2000.

[23] C. Xu and K. Kwak, "Comments on closed-form analysis of linearly constrained CMA-based blind multiuser detector," *IEEE Commun. Lett.*, July 2001.

[24] X. Li, "Linear prediction approach for joint blind equalization and multiuser detection in CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, pp. 3134–3145, Nov. 2000.

[25] I. Ghauri and D. T. M. Slock, "Blind and semi-blind single user receiver techniques for asynchronous CDMA in multipath channels," *Proc. Globecom, Sydney, Australia*, vol. 1998, Nov. 1998.

[26] X. Li and H. Fan, "Direct estimation of blind zero-forcing equalizers based on second order statistics," *IEEE Transactions on Signal Processing*, vol. 48, pp. 2211–2218, Aug. 2000.

[27] J. G. Proakis, *Digital Communications*. McGraw-Hill, 3rd ed., 1995.

[28] J. K. Tungait, L. Tong, and Z. Ding, "Recent research in training based and blind approaches," *IEEE Signal Processing Magazine*, pp. 17–28, May 2000.

[29] J. P. de Villiers and L. P. Linde, "On the convexity of the LCCM cost function for DS-CDMA blind multiuser detection," *Submitted, not yet approved - IEEE Trans. Commun.*, June 2002.

[30] J. P. de Villiers and L. P. Linde, "Analysis of the LCDCM criterion for DS-CDMA blind multiuser detection," *Proceedings of the IEEE Africon*, October 2002.

[31] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.

[32] T. S. Rappaport, *Wireless Communications*. New Jersey: Prentice Hall, 1996.

[33] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Systems Technical Journal*, vol. 47, pp. 957–1000, 1968.

[34] H. V. Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

[35] L. J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics*. Belmont, California: Duxbury Press, 2nd ed., 1992.

[36] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–12, 1949.

[37] J. Wozencroft and I. Jacobs, *Priciples of Communications Engineering*. New York: Wiley, 1965.

[38] M. Pursley, "Performance evaluation for phase-coded spread spectrum multiple-access communication - part i: System analysis," *IEEE Trans. Commun.*, vol. COM-25, pp. 795–799, Aug. 1977.

[39] K. Yao, "Error probability of asynchronous spread spectrum multiple access communication systems," *IEEE Trans. Commun.*, vol. COM-25, pp. 803–809, Aug. 1977.

[40] S. Verdu and S. Shamai, "Multiuser detection with random spreading and error correction codes: Fundamental limits.," in *Proc. 35th Allerton Conf. Communications, Control and Computing*, Sept. - Oct. 1997.

[41] B. Saltzberg, "Intersymbol interference error bounds with application to ideal bandlimited signalling," *IEEE Trans. Information Theory*, vol. 14, pp. 563–568, July 1968.

[42] J. S. Lehnert and M. B. Pursley, "Error probabilities for binary direct-sequence spread-spectrum communications with random signature sequences," *IEEE Trans. Commun.*, vol. COM-35, pp. 87–98, Jan. 1987.

[43] J. M. Holtzman, "A simple, accurate method to calculate spread-spectrum multiple-access error probabilities," *IEEE Trans. Commun.*, vol. 40, pp. 461–464, March 1992.

[44] P. van Rooyen and F. Solms, "Maximum entropy investigation of inter user interference distribution in a DS/SSMA system," in *Proc. 1995 IEEE Personal, Indoor, Mobile Radio Communications Conf.*, pp. 1308–1312, Sept. 1995.

[45] S. Verdu, "Optimum multiuser asymptotic efficiency," *IEEE Trans. Commun.*, vol. COM-34, pp. 890–897, Sept. 1986.

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

[47] N. R. Mangalvedhe, *Development and Analysis of Adaptive Intereference Rejection Techniques for Direct sequence Code Division Multiple Access Systems*. PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, July 1999.

[48] K. S. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Trans. Aerosp. Electron. Syst.,*, vol. AES-15, pp. 181–185, Jan. 1979.

[49] T. Meyers and M. E. Magana, "An adaptive implementation of the 'one shot' decorrelating detector for CDMA communications," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 762–765, Sept. 1997.

[50] D. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE Journal on Selected Areas in Commun.*, vol. 12, pp. 808–816, June 1994.

[51] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[52] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE Journal on Selected Areas in Commun.*, vol. 12, pp. 685–697, May 1994.

[53] S. L. Miller, "Training analysis of adaptive interference suppression for direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 488–495, April 1996.

[54] P. Monogioudis, R. Tafazolli, and B. Evans, "Performance of adaptive nonlinear NEFAR CDMA receiver architecture," *Electronics Letters*, vol. 30, Feb. 1994.

[55] L. B. Milstein, "Interference rejection techniques in spread spectrum communications," *Proc. IEEE*, vol. 76, pp. 657–671, June 1988.

[56] N. J. Bershad, "Error probabilities for DS spread-spectrum systems using an ALE for narrow-band interference rejection," *IEEE Trans. Commun.*, vol. 36, pp. 588–595, May 1988.

[57] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice Hall, 1991.

[58] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. New Jersey: Prentice-Hall, 1985.

[59] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. on Info. Theory*, vol. 43, pp. 858–871, May 1997.

[60] T.Miyajima, "Blind adaptive detection using differential CMA for CDMA systems," *IEICE Trans.A*, vol. J83-A, pp. 1318–1329, Nov. 2000.

[61] O. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[62] D. Goddard, "Self recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

[63] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459–472, April 1983.

[64] W. Lee, B. Vojcic, and R. Pickholz, "Constant modulus algorithm for blind multiuser detection," in *Proc. ISSSTA*, (Mainz, Germany), pp. 1262–1266, Sept. 1996.

[65] F. A. Valentine, *Convex sets*. Mcgraw Hill, 1964.

[66] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, 1979.

[67] L. Castedo and A. R. Figueiras-Vidal, "An adaptive beamforming technique based on cyclostationary signal properties," *IEEE Trans. Signal Processing*, vol. 43, pp. 1637–1650, July 1995.

[68] X. Wang and H. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, pp. 677–690, March 1998.

[69] L. P. Linde, F. Marx, and W. Malan, "Power and spectral efficiency of a family of constant envelope root-of-unity filtered complex spreading sequences in WCDMA non-linear power amplification," in *IEEE Africon*, October 2002.

[70] L. Schwartz, *Cours d'Analyse*, vol. 2. Paris: Hermann, 1967. pp. 271-278.

[71] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application*. Addison-Wesley, 1974.

APPENDIX A

SELECTED PROPERTIES OF THE Q-FUNCTION

This appendix contains a number of properties of the Q function that are used in the text of this dissertation.

1.
$$Q(x) = \int_x^\infty \frac{1}{2\pi} e^{-t^2/2} dt. \tag{A.1}$$

2.
$$Q(x) = P[X > x], \tag{A.2}$$

where X is a zero-mean, unit-variance Gaussian random variable.

3. $Q(x)$ is monotonically decreasing.

4.
$$\lim_{\sigma \to 0} \frac{Q\left(\frac{\alpha}{\sigma}\right)}{Q\left(\frac{\beta}{\sigma}\right)} = \begin{cases} +\infty, & [\alpha]^+ < \beta; \\ 2, & \alpha < \beta = 0; \\ 1, & \alpha = \beta \text{ or } \max\{\alpha, \beta\} < 0; \\ 1/2, & \beta < \alpha = 0; \\ 0, & [\beta]^+ < \alpha; \end{cases} \tag{A.3}$$

where

$$[z]^+ = \max\{0, z\}. \tag{A.4}$$

5.
$$2 \lim_{\sigma \to 0} \sigma^2 \log Q\left(\frac{x}{\sigma}\right) = -\left([x]^+\right)^2. \tag{A.5}$$

6. If X is a zero-mean, unit variance, normal random variable, then

$$E\left[Q\left(\mu+\lambda X\right)\right]=Q\left(\frac{\mu}{\sqrt{1+\lambda^2}}\right). \tag{A.6}$$

7.

$$\int_0^\infty x\exp\left(-\frac{x^2}{2}\right)Q\left(\frac{x}{\sigma}\right)dx=\frac{1}{2}\left(1-\frac{1}{\sqrt{1+\sigma^2}}\right). \tag{A.7}$$

APPENDIX B

SIMULATION OF MOBILE CHANNEL

This appendix contains the mobile channel model used to obtain the simulation results. In the first section, the Doppler filter is described, and how it is used in a baseband Rayleigh fading simulator. In the second section, the Rayleigh fading simulator in the first section is used to implement several independent fading paths in a mobile fading simulator.

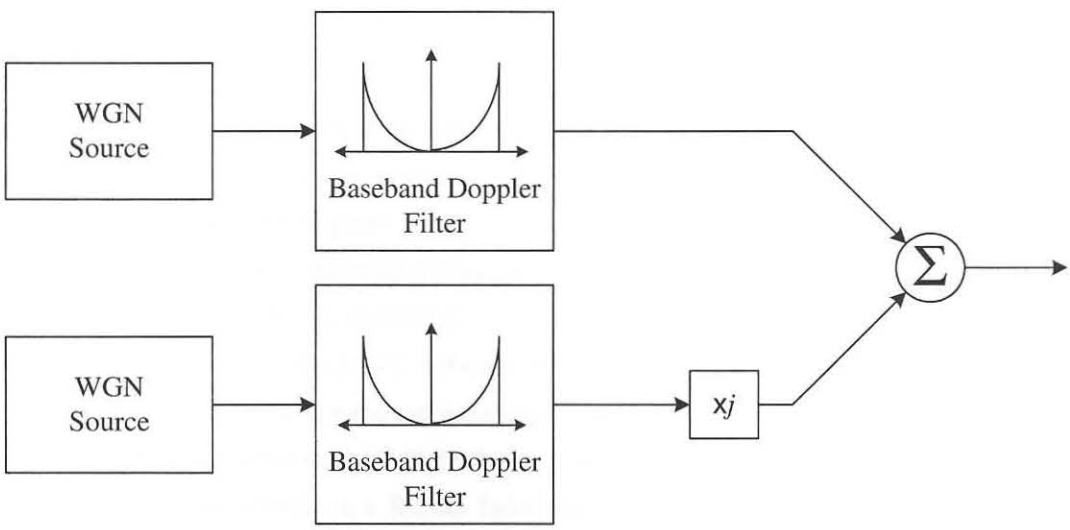

Figure B.2: Baseband complex Rayleigh fading coefficient simulator.

B.1 DOPPLER SPREAD RAYLEIGH FADING

To be able to simulate Clarke's model, we will have to analyze the spectrum of received electric field. This is derived in [32] and is shown to be

$$S_{E_z}(f) = \frac{1.5}{\pi f_m \sqrt{1 - \left(\frac{f - f_c}{f_m}\right)}} \tag{B.1}$$

with a vertical $\lambda/4$ antenna ($G(\alpha) = 1.5$), and a uniform distribution of incoming power over 0 to 2π. In this equation, f_m is maximum Doppler shift and f_c the carrier frequency. The baseband power spectral density is given by equation (B.2),

$$S_{bbE_z}(f) = \frac{1}{8\pi f_m} K\left[\sqrt{1 - \left(\frac{f}{2f_m}\right)^2}\right], \tag{B.2}$$

where $K[\cdot]$ is the complete elliptical integral of the first kind. When we wish to simulate the Rayleigh fading channel, we can do this by sending both the in-phase and quadrature baseband independent Gaussian noise samples through baseband filters with the transfer function given in equation (B.2). The resulting complex signal can then be utilized as a complex Rayleigh fading coefficient in a baseband simulation environment.

In this dissertation, a infinite impulse response (IIR) third order approximation of a 50Hz Doppler filter is used for simulation purposes. The power spectral density of this filter is shown in Figure B.3.

B.2 FREQUENCY SELECTIVE MULTIPATH RAYLEIGH FADING

In the mobile channel, the receiver antenna picks up the sum of independent Rayleigh (Doppler spread) faded multiple paths. In the previous section, the simulator that implements the fading coefficients of each of these paths is given. Here we present the simulation model to implement multiple independent fading paths with different delays. This introduces the harsh frequency selectivity that is frequently encountered in mobile channel environments. The frequency selective (multipath) mobile channel model for simulation purposes is shown in Figure B.4. This is basically a finite impulse response (FIR) filter structure, where the tap weights are the path strengths corresponding to each of the time delay- or multipath components. Note that to get a line-of-sight component, an unfaded signal may be added. This will result in a Ricean faded signal. It is possible to vary the strength of the specular (line-of-sight) path by a constant weighting of this component.

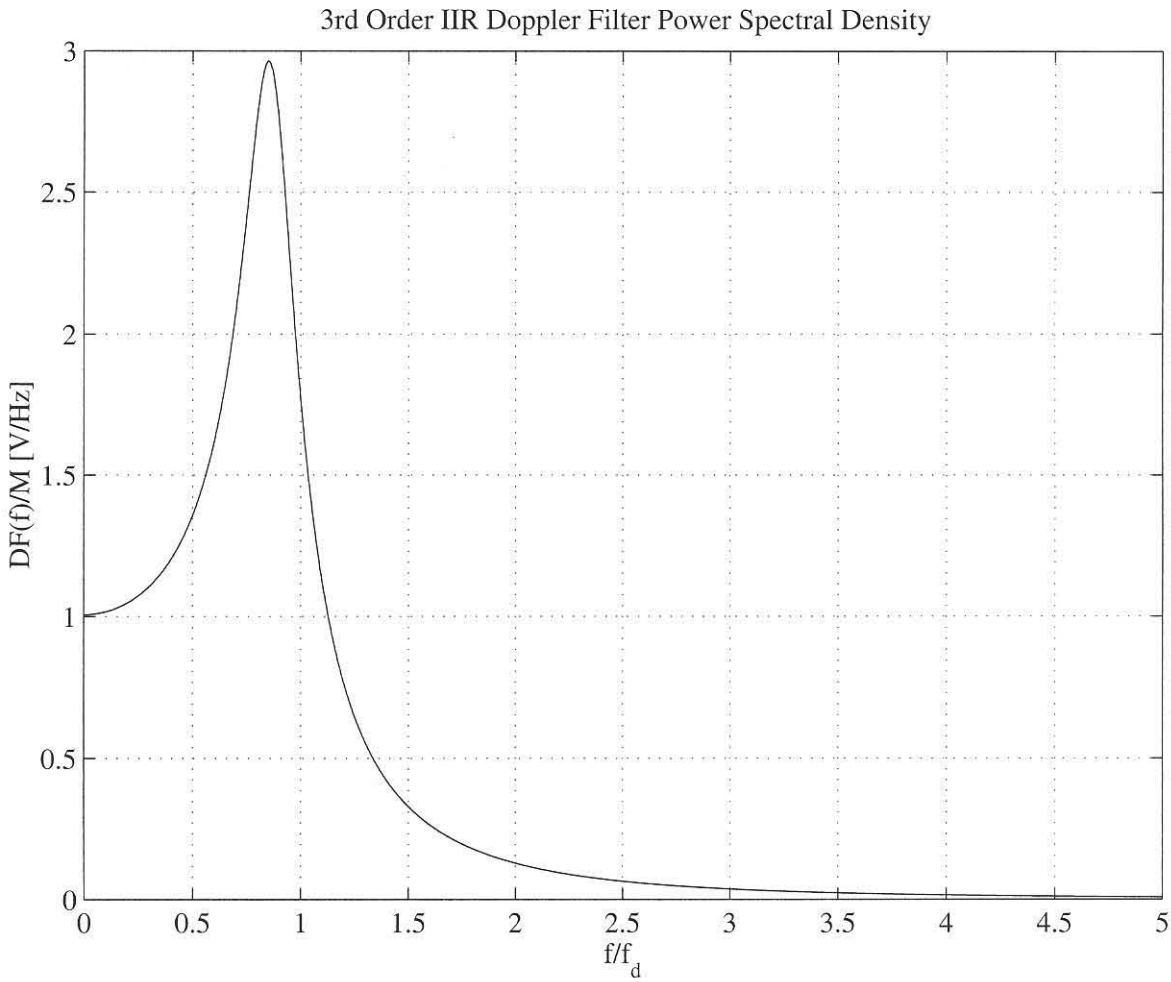

Figure B.3: Frequency spectrum of 3rd order approximation of a Doppler filter with a Doppler frequency of 50Hz.

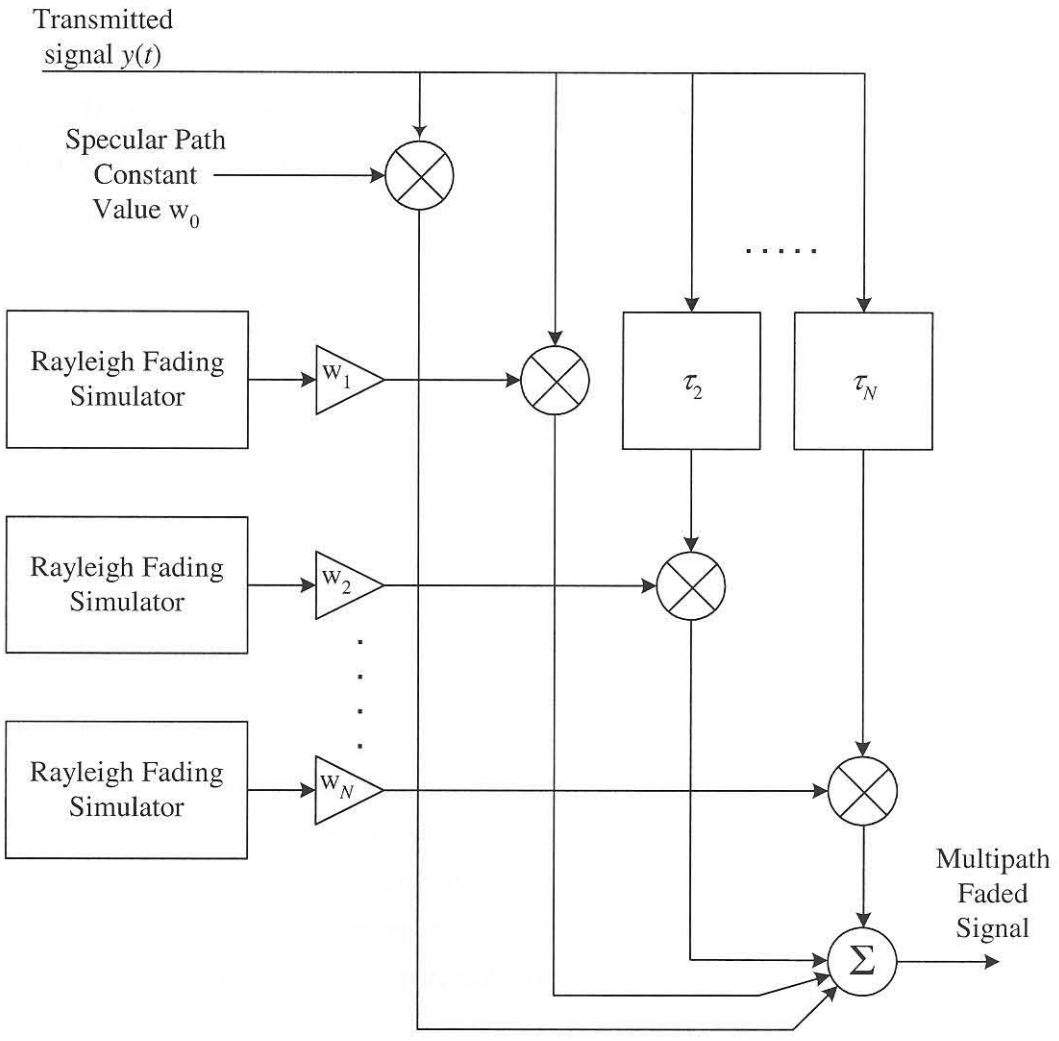

Figure B.4: Model of a frequency selective (multipath) fading channel.

APPENDIX C

DIFFERENTIATION WITH RESPECT TO A COMPLEX MATRIX

An issue commonly encountered in the study of optimization theory, is that of differentiating a cost function with respect to a parameter vector or matrix. In the text, the normal gradient operator is used. The differentiation of a cost function with respect to a complex vector or matrix is more involved. This appendix will expand on the case of differentiation with respect to a complex vector in Appendix B of [57], to the case of differentiation with respect to a complex matrix. The relationship between the concepts of a gradient and a derivative for complex matrices is discussed here.

C.1 BASIC DEFINITIONS

Consider a complex function $f(\mathbf{M})$ that is dependent on a parameter matrix $\mathbf{M}$. When the entries of $\mathbf{M}$ are complex valued, there are two different mathematical concepts that require individual attention: (1) the matrix nature of $\mathbf{M}$, and (2) the fact that each entry of $\mathbf{M}$ is complex valued.
Let us start with the fact that the element of the of the lth row and the mth column of the matrix $\mathbf{M}$ is the sum of a real part and an imaginary part multiplied by $j = \sqrt{-1}$, i.e.

$$m_{vw} = x_{vw} + jy_{vw}. \tag{C.1}$$

The real and imaginary parts of equation (C.1) can alternatively be written in terms of the pair of complex conjugates m_{vw} and m_{vw}^* with

$$x_{vw} = \frac{1}{2}\left(m_{vw} + m_{vw}^*\right) \tag{C.2}$$

and

$$y_{vw} = \frac{1}{2j}\left(m_{vw} - m_{vw}^*\right), \tag{C.3}$$

where $(\cdot)^*$ denotes complex conjugation. It is evident that the real quantities x_{vw} and y_{vw} are functions of both m_{vw} and m^*_{vw}. Only when we deal with an analytic function f, may we abandon the complex-conjugated term m^*_{vw} by virtue of the Cauchy-Riemann equations. However, it is rare that one encounters analytic functions in physical sciences and engineering.

When considering a derivative, a connection should be made with the concept of a differential. In particular, the chain rule of changes of variables must be satisfied. Considering these points, the mathematician Schwartz [70] defined certain complex derivatives in terms of real derivatives:

$$\frac{\partial}{\partial m_{vw}} = \frac{1}{2}\left(\frac{\partial}{\partial x_{vw}} - j\frac{\partial}{\partial y_{vw}}\right) \tag{C.4}$$

and

$$\frac{\partial}{\partial m^*_{vw}} = \frac{1}{2}\left(\frac{\partial}{\partial x_{vw}} + j\frac{\partial}{\partial y_{vw}}\right). \tag{C.5}$$

The above derivatives satisfy the following basic requirements with respect to a differential:

$$\frac{\partial m_{vw}}{\partial m_{vw}} = 1 \tag{C.6}$$

and

$$\frac{\partial m_{vw}}{\partial m^*_{vw}} = \frac{\partial m^*_{vw}}{\partial m_{vw}} = 0. \tag{C.7}$$

(An analytic function f satisfies $\frac{\partial f}{\partial z^*} = 0$ everywhere.) Equations (C.4) and (C.5) are referred to as the *derivative* and the *conjugate derivative* respectively, both with respect to m_{vw}.

Let us extend this notion to the general case of the derivative with respect to a matrix with complex elements. Extension of the derivative with respect to a complex vector was done in [71] and was also dealt with in Appendix B of [57]. We can extend equations (C.4) and (C.5) to the derivative with respect to a $K \times K$ matrix:

$$\frac{\partial}{\partial \mathbf{M}} = \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial x_{11}} - j\frac{\partial}{\partial y_{11}} & \frac{\partial}{\partial x_{12}} - j\frac{\partial}{\partial y_{12}} & \cdots & \frac{\partial}{\partial x_{1K}} - j\frac{\partial}{\partial y_{1K}} \\ \frac{\partial}{\partial x_{21}} - j\frac{\partial}{\partial y_{21}} & \frac{\partial}{\partial x_{22}} - j\frac{\partial}{\partial y_{22}} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_{K1}} - j\frac{\partial}{\partial y_{K1}} & \cdots & \cdots & \frac{\partial}{\partial x_{KK}} - j\frac{\partial}{\partial y_{KK}} \end{bmatrix} \tag{C.8}$$

and

$$\frac{\partial}{\partial \mathrm{M}^*} = \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial x_{11}} + j\frac{\partial}{\partial y_{11}} & \frac{\partial}{\partial x_{12}} + j\frac{\partial}{\partial y_{12}} & \cdots & \frac{\partial}{\partial x_{1K}} + j\frac{\partial}{\partial y_{1K}} \\ \frac{\partial}{\partial x_{21}} + j\frac{\partial}{\partial y_{21}} & \frac{\partial}{\partial x_{22}} + j\frac{\partial}{\partial y_{22}} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_{K1}} + j\frac{\partial}{\partial y_{K1}} & \cdots & \cdots & \frac{\partial}{\partial x_{KK}} + j\frac{\partial}{\partial y_{KK}} \end{bmatrix}. \tag{C.9}$$

Analogous to the scalar case, the above two derivatives obey the following relations:

$$\frac{\partial \mathrm{M}}{\partial \mathrm{M}} = \mathbf{I} \tag{C.10}$$

and

$$\frac{\partial \mathrm{M}}{\partial \mathrm{M}^*} = \frac{\partial \mathrm{M}^*}{\partial \mathrm{M}} = \mathbf{0} \tag{C.11}$$

where $\mathbf{I}$ and $\mathbf{0}$ are the $K^2 \times K^2$ identity and null matrices respectively.

For the purpose of differentiating with respect to complex matrices, equation (C.9) (the *conjugate derivative*) will be adopted as the *derivative with respect to a complex valued matrix*.

C.2 THE GRADIENT MATRIX IN TERMS OF THE DERIVATIVE WITH RESPECT TO A MATRIX

Consider a real cost function $J(\mathrm{M})$ that defines the $K \times K$ dimension error performance surface in terms of the $K \times K$ matrix M. The *complex gradient matrix* was defined in Chapter 4 as

$$\nabla_{\mathrm{M}}(J) = \begin{bmatrix} \frac{\partial J}{\partial x_{11}} + j\frac{\partial J}{\partial y_{11}} & \cdots & \frac{\partial J}{\partial x_{1K}} + j\frac{\partial J}{\partial y_{1K}} \\ \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial x_{K1}} + j\frac{\partial J}{\partial y_{K1}} & \cdots & \frac{\partial J}{\partial x_{KK}} + j\frac{\partial J}{\partial y_{KK}} \end{bmatrix}, \tag{C.12}$$

where $x_{vw} + jy_{vw}$ is the element of the vth row and the wth column of the linear transformation M. The gradient matrix is *normal* to the $K \times K$ dimensional error surface. If we relate (C.12) to (C.9), we find that the gradient is related to the conjugate derivative by

$$\nabla_{\mathrm{M}}(J) = 2\frac{\partial J}{\partial \mathrm{M}^*}. \tag{C.13}$$

This means that the gradient is a scaled form of the conjugate derivative of (C.9).

C.3 DIFFERENTIATING THE COMPONENTS OF THE MMSE COST FUNCTION

With these preceding ideas in mind, let us attempt to differentiate the cost function in (4.29) to arrive at the result in (4.31). Let us start with the gradient of the cost function:

$$4\frac{\partial}{\partial \mathbf{M}^*}\left(\text{tr}\left\{\mathbf{I}\right\} - \text{tr}\left\{\mathbf{ARM}^H\right\} - \text{tr}\left\{\mathbf{MAR}\right\} + \text{tr}\left\{\mathbf{MRA}^2\mathbf{RM}^H\right\} + \text{tr}\left\{\mathbf{M}\sigma^2\mathbf{RM}^H\right\}\right). \tag{C.14}$$

Since differentiation is a linear operation, we can differentiate the terms of the cost function individually. It is trivial to see that the differentiation of the first term results in zero, i.e.

$$\frac{\partial}{\partial \mathbf{M}^*}\text{tr}\left\{\mathbf{I}\right\} = 0. \tag{C.15}$$

To find the derivative of the second term of (C.14), we will exploit the definition of the trace of a matrix. We start with the ijth element of the product $\mathbf{ARM}^H$. We have

$$\left(\mathbf{ARM}^H\right)_{ij} = \sum_l a_{il}(\mathbf{RM})_{lj} = \sum_l \sum_k a_{il} r_{lk} m^*_{jk}, \tag{C.16}$$

where a_{il}, r_{lk} and m^*_{jk} are the complex elements of the matrices $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{M}^H$ respectively. The diagonal elements of the product $\mathbf{ARM}^H$ are

$$\left(\mathbf{ARM}^H\right)_{ii} = \sum_l \sum_k a_{il} r_{lk} m^*_{ik}. \tag{C.17}$$

Since the trace of a matrix (or product of matrices) is the sum of the diagonal elements, we have

$$\text{tr}\left\{\mathbf{ARM}^H\right\} = \sum_i \sum_l \sum_k a_{il} r_{lk} m^*_{ik}. \tag{C.18}$$

To apply $\frac{\partial}{\partial \mathbf{M}^*}$ to the trace of $\mathbf{ARM}^H$, we turn our attention to the derivative with respect to the individual elements $\frac{\partial}{\partial m^*_{ik}}$. Applying this to (C.18), we obtain

$$\begin{aligned}
\frac{\partial}{\partial m^*_{ik}}\left(\text{tr}\left\{\mathbf{ARM}^H\right\}\right) &= \frac{\partial}{\partial m^*_{ik}}\left(\sum_i \sum_l \sum_k a_{il} r_{lk} m^*_{ik}\right) \\
&= \sum_l a_{il} r_{lk} \qquad \text{(C.19)} \\
&= \left(\mathbf{AR}\right)_{ik}, \qquad \text{(C.20)}
\end{aligned}$$

and consequently

$$\frac{\partial}{\partial \mathbf{M}^*}\left(\text{tr}\left\{\mathbf{ARM}^H\right\}\right) = \mathbf{AR}. \tag{C.21}$$

To determine the complex derivative of the third term, $\frac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\left\{\mathbf{MAR}\right\}\right)$, we follow similar reasoning. In this case we have

$$\mathrm{tr}\left\{\mathbf{MAR}\right\} = \sum_i \sum_l \sum_k m_{il} a_{lk} r_{ki}. \tag{C.22}$$

Again, applying the derivative with respect to the individual elements $\frac{\partial}{\partial m_{ik}^*}$ we get

$$\begin{aligned} \frac{\partial}{\partial m_{il}^*}\left(\mathrm{tr}\left\{\mathbf{MAR}\right\}\right) &= \frac{\partial}{\partial m_{il}^*}\left(\sum_i \sum_l \sum_k m_{il} a_{lk} r_{ki}\right) \\ &= 0, \end{aligned} \tag{C.23}$$

since $\frac{\partial m_{il}}{\partial m_{il}^*} = 0$ from (C.7). Thus we have

$$\frac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\left\{\mathbf{MAR}\right\}\right) = \mathbf{0}, \tag{C.24}$$

where $\mathbf{0}$ is the null matrix.

The derivatives of quadratic terms in (C.14) can be evaluated by means of the product rule and following the same reasoning as with the previous two terms. Let us first evaluate

$$\frac{\partial}{\partial \mathbf{M}^*}\mathrm{tr}\left\{\mathbf{MRA}^2\mathbf{RM}^H\right\}. \tag{C.25}$$

To simplify matters, we let $\mathbf{B} = \mathbf{RA}^2\mathbf{R}$. This means that

$$\mathrm{tr}\left\{\mathbf{MBM}^H\right\} = \sum_i \sum_l \sum_k m_{il} b_{lk} m_{ik}^*. \tag{C.26}$$

Differentiating (C.26) by means of the product rule, we obtain

$$\begin{aligned} \frac{\partial}{\partial m_{il}^*}\left(\mathrm{tr}\left\{\mathbf{MBM}^H\right\}\right) &= \frac{\partial}{\partial m_{il}^*}\left(\sum_i \sum_l \sum_k m_{il} b_{lk} m_{ki}^*\right) \\ &= 0 + \sum_l m_{il} b_{lk} \qquad \text{(C.27)} \\ &= (\mathbf{MB})_{ik} \qquad \text{(C.28)} \end{aligned}$$

where (C.27) follows from (C.6) and (C.7). Thus we have

$$\frac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\left\{\mathbf{MBM}^H\right\}\right) = \mathbf{MB}. \tag{C.29}$$

If we apply this result to the two quadratic terms, we obtain

$$\frac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\left\{\mathbf{MRA}^2\mathbf{RM}^H\right\}\right) = \mathbf{MRA}^2\mathbf{R} \tag{C.30}$$

and

$$\frac{\partial}{\partial \mathbf{M}^*}\left(\mathrm{tr}\left\{\mathbf{M}\sigma^2\mathbf{R}\mathbf{M}^H\right\}\right) = \mathbf{M}\sigma^2\mathbf{R}. \tag{C.31}$$

Using all of the above results, the gradient of the cost function J is given by

$$\nabla_{\mathbf{M}}(J) = -4\mathbf{A}\mathbf{R} + 4\mathbf{M}\mathbf{R}\mathbf{A}^2\mathbf{R} + 4\mathbf{M}\sigma^2\mathbf{R} \tag{C.32}$$

Thus verifying eq. (4.31).

APPENDIX D

EVALUATING THE EXPECTED VALUE IN THE LCCM AND LCDCM COST FUNCTIONS

D.1 EXPECTED VALUE IN THE LCCM COST FUNCTION

From the first term in equation (5.3) we have

$$\begin{aligned} E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] &= E\left[\left(\sum_{i=1}^{K}\sum_{j=1}^{K}u_i^*b_ib_j^*u_j\right)\left(\sum_{k=1}^{K}\sum_{l=1}^{K}u_k^*b_kb_l^*u_l\right)\right] \\ &= E\left[\sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K}u_i^*b_ib_j^*u_ju_k^*b_kb_l^*u_l\right] \\ &= E\left[\sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K}u_i^*u_ju_k^*u_lb_ib_j^*b_kb_l^*\right] \\ &= \sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(u_i^*u_ju_k^*u_lE\left[b_ib_j^*b_kb_l^*\right]\right). \end{aligned} \tag{D.1}$$

Consider equation (D.1). Since the different users' bits are independent, a zero result will be produced if $i \neq j \neq k \neq l$. If $i = k$ and $l = j$, and $b_k \in \{\pm 1 \pm j\}$ the argument of the expected value produces a result of either 4 or -4, each with a probability of $P(4) = P(-4) = 0.5$. It is easy to see that the expected value in this case will also be zero. There are three cases in which the expected value will produce a nonzero result. These are:

1. $i = j = k = l$;
2. $i = j$ and $k = l$, but $i = j \neq k = l$;
3. $j = k$ and $i = l$, but $j = k \neq i = l$.

Taking all of the above into account, we can write (D.1) as

$$\begin{aligned} E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] &= 4\left(\sum_{k=1}^{K} u_k^* u_k u_k^* u_k + \sum_{i=1}^{K}\sum_{\substack{k=1\\ i\neq k}}^{K} u_i^* u_i u_k^* u_k + \sum_{j=1}^{K}\sum_{\substack{l=1\\ j\neq l}}^{K} u_i^* u_i u_k^* u_k\right) \\ &= 4\left(\sum_{i=1}^{K}\sum_{k=1}^{K} u_i^* u_i u_k^* u_k + \sum_{j=1}^{K}\sum_{\substack{l=1\\ j\neq l}}^{K} u_i^* u_i u_k^* u_k\right) \\ &= 4\left(\sum_{i=1}^{K}\sum_{k=1}^{K} u_i^* u_i u_k^* u_k + \sum_{j=1}^{K}\sum_{l=1}^{K} u_i^* u_i u_k^* u_k - \sum_{k=1}^{K} u_k^* u_k u_k^* u_k\right) \\ &= 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4. \end{aligned} \tag{D.2}$$

In the same way we can write the expectation from the second term in equation (5.3) as

$$\begin{aligned} E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] &= E\left[\sum_{i=1}^{K}\sum_{j=1}^{K} u_i^* b_i b_j^* u_j\right] \\ &= \sum_{i=1}^{K}\sum_{j=1}^{K} u_i^* u_j E\left[b_i b_j^*\right]. \end{aligned} \tag{D.3}$$

The expected value will produce a nonzero result only if $i = j$. In this case we have

$$\begin{aligned} E\left[\mathbf{u}^H\mathbf{b}\mathbf{b}^H\mathbf{u}\right] &= 2\sum_{i=1}^{K}\sum_{j=1}^{K} u_i^* u_j \\ &= 2\mathbf{u}^H\mathbf{u}. \end{aligned} \tag{D.4}$$

D.2 EXPECTED VALUE IN THE LCDCM COST FUNCTION

Exactly in the same way as we have derived (D.2), we can show that the first and last terms of equation (5.26) are

$$E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\right] = 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_k|^4, \tag{D.5}$$

and

$$E\left[\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] = 8\left(\mathbf{u}^H\mathbf{u}\right)^2 - 4\sum_{k=1}^{K}|u_{k.}|^4. \tag{D.6}$$

Since D is large enough for bits separated by D seconds from the same user to be independent, we have from (5.26) the middle term as

$$E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] = \sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(u_i^* u_j u_k^* u_l E\left[b_i[i]b_j^*[i]b_k[i-D]b_l^*[i-D]\right]\right). \tag{D.7}$$

In this case only two possibilities will produce a nonzero expected value result. These are:

1. $i = j = k = l$;
2. $i = j$ and $k = l$, but $i = j \neq k = l$.

Using this, we have equation (D.7) equal to

$$\begin{aligned} E\left[\mathbf{u}^H\mathbf{b}[i]\mathbf{b}^H[i]\mathbf{u}\mathbf{u}^H\mathbf{b}[i-D]\mathbf{b}^H[i-D]\mathbf{u}\right] &= 4\left(\sum_{k=1}^{K} u_k^* u_k u_k^* u_k + \sum_{i=1}^{K}\sum_{\substack{k=1\\ i\neq k}}^{K} u_i^* u_i u_k^* u_k\right) \\ &= 4\left(\sum_{i=1}^{K}\sum_{k=1}^{K} u_i^* u_i u_k^* u_k\right) \\ &= 4\left(\mathbf{u}^H\mathbf{u}\right)^2. \end{aligned} \tag{D.8}$$