1

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395 rue Wellington
Ottawa (Ontario)
K1A 0N4

395 - 396 - 397 - 398 - 399

395 - 396 - 397 - 398 - 399

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Co-channel and Adjacent-channel Interference Cancellation Using Neural Networks

by

Xianrong Yao

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of

Master of Engineering

Ottawa-Carleton Institute for Electrical Engineering
Faculty of Engineering
Department of Systems and Computer Engineering
Carleton University
Ottawa, Ontario, Canada, K1S 5B6

April 10, 1993

© Xianrong Yao

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-89899-2

Canada

Name _____

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

--	--	--	--

U·M·I

SUBJECT TERM

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION

General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language
 General 0679
 Ancient 0289
 Linguistics 0290
 Modern 0291
Literature
 General 0401
 Classical 0294
 Comparative 0295
 Medieval 0297
 Modern 0298
 African 0316
 American 0591
 Asian 0305
 Canadian (English) 0352
 Canadian (French) 0355
 English 0593
 Germanic 0311
 Latin American 0312
 Middle Eastern 0315
 Romance 0313
 Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion
 General 0318
 Biblical Studies 0321
 Clergy 0319
 History of 0320
 Philosophy of 0322
Theology 0469

SOCIAL SCIENCES

American Studies 0323
Anthropology
 Archaeology 0324
 Cultural 0326
 Physical 0327
Business Administration
 General 0310
 Accounting 0272
 Banking 0770
 Management 0454
 Marketing 0338
Canadian Studies 0385
Economics
 General 0501
 Agricultural 0503
 Commerce-Business 0505
 Finance 0508
 History 0509
 Labor 0510
 Theory 0511
Folklore 0358
Geography 0366
Gerontology 0351
History
 General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0392
Political Science
 General 0615
 International Law and Relations 0616
 Public Administration 0617
Recreation 0814
Social Work 0452
Sociology
 General 0626
 Criminology and Penology 0627
 Demography 0938
 Ethnic and Racial Studies 0631
 Individual and Family Studies 0628
 Industrial and Labor Relations 0629
 Public and Social Welfare 0630
 Social Structure and Development 0700
 Theory and Methods 0344
Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture
 General 0473
 Agronomy 0285
 Animal Culture and Nutrition 0475
 Animal Pathology 0476
 Food Science and Technology 0359
 Forestry and Wildlife 0478
 Plant Culture 0479
 Plant Pathology 0480
 Plant Physiology 0817
 Range Management 0777
 Wood Technology 0746
Biology
 General 0306
 Anatomy 0287
 Biostatistics 0308
 Botany 0309
 Cell 0772
 Ecology 0329
 Entomology 0353
 Genetics 0369
 Limnology 0793
 Microbiology 0410
 Molecular 0307
 Neuroscience 0317
 Oceanography 0416
 Physiology 0433
 Radiation 0821
 Veterinary Science 0778
 Zoology 0472
Biophysics
 General 0786
 Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Polynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
Health Sciences
 General 0566
 Audiology 0300
 Chemotherapy 0992
 Dentistry 0567
 Education 0350
 Hospital Management 0769
 Human Development 0758
 Immunology 0982
 Medicine and Surgery 0564
 Mental Health 0547
 Nursing 0569
 Nutrition 0570
 Obstetrics and Gynecology 0380
 Occupational Health and Therapy 0354
 Ophthalmology 0381
 Pathology 0571
 Pharmacology 0419
 Pharmacy 0572
 Physical Therapy 0382
 Public Health 0573
 Radiology 0574
 Recreation 0575

Speech Pathology 0460
Toxicology 0363
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
Chemistry
 General 0485
 Agricultural 0749
 Analytical 0486
 Biochemistry 0487
 Inorganic 0488
 Nuclear 0738
 Organic 0490
 Pharmaceutical 0491
 Physical 0494
 Polymer 0495
 Radiation 0754
Mathematics 0405
Physics
 General 0605
 Acoustics 0986
 Astronomy and Astrophysics 0606
 Atmospheric Science 0608
 Atomic 0748
 Electronics and Electricity 0607
 Elementary Particles and High Energy 0798
 Fluid and Plasma 0759
 Molecular 0609
 Nuclear 0610
 Optics 0752
 Radiation 0756
 Solid State 0611
Statistics 0463
Applied Sciences
Applied Mechanics 0346
Computer Science 0984

Engineering
 General 0537
 Aerospace 0538
 Agricultural 0539
 Automotive 0540
 Biomedical 0541
 Chemical 0542
 Civil 0543
 Electronics and Electrical 0544
 Heat and Thermodynamics 0348
 Hydraulic 0545
 Industrial 0546
 Marine 0547
 Materials Science 0794
 Mechanical 0548
 Metallurgy 0743
 Mining 0551
 Nuclear 0552
 Packaging 0549
 Petroleum 0765
 Sanitary and Municipal 0554
 System Science 0790
Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

The undersigned recommend to the Faculty of
Graduate Studies and Research acceptance of the thesis

**“Co-channel and Adjacent-channel Interference
Cancellation Using Neural Networks”**

submitted by
Xianrong Yao
in partial fulfillment of the requirements
for the degree of Master of Engineering

Thesis Supervisor

Chair,
Department of Systems and Computer Engineering

Department of Systems and Computer Engineering
Faculty of Engineering
Carleton University
May 11, 1993

Abstract

This thesis investigates the problem of adaptive equalization in the presence of intersymbol interference, additive white Gaussian noise, co-channel and adjacent-channel interference. The multilayer perceptron-based decision feedback equalizer (MLPDFE) is designed to realize a sophisticated non-linear adaptive equalizer capable of operating under poor signal to interference ratio, poor signal to noise ratio and server timing misalignment of co-channels with respect to the main channel conditions.

This thesis starts by a demonstration of the sensitivity of the MLPDFE performance (in terms of averaged MSE) on its perceptron configuration and adaptive parameters. The effort of co-channel and adjacent-channel interference on the system performance is investigated separately. The results of a computer simulation study to compare the performance of the MLPDFE with that of the least mean square decision feedback equalizer (LMSDFE) in terms of the SER with a QPSK signalling are presented. The results indicate that the MLPDFE provides better SER performance relative to the LMSDFE when only co-channel interference presents, especially in server co-channel timing misalignment conditions. Also, in the presence of only adjacent-channel interference, the MLPDFE has very similar SER performance as that of the LMSDFE.

Acknowledgments

I wish to thank my thesis supervisor, Dr. H. M. Hafez, for his consistent guidance and encouragement during the course of this thesis research. My deepest gratitude is due to him for his advice, patience, and understanding.

I would also like to thank my fellow former and current graduates in particular, Shahram Golastaneh, Pimpin Kwala and Norm Lo for their help and many enlightening discussions and suggestions. Furthermore, I like to thank the staff members in the System and Computer Engineering Department for their continuing support.

*To :
My parents,
for their love and support*

Table of Contents

Abstract	iii
Acknowledgments	iv
List of Figures	ix
List of Tables	xii
List of Abbreviations and Acronyms	xiii
Lists of Mathematical Notation	xv
Chapter 1 Introduction	1
1.1 Problem Definition	2
1.2 Thesis Objectives	3
1.3 Thesis Organization	3
Chapter 2 Background Information and Literature Survey	5
2.1 An Overview of Neural Networks	5
2.1.1 Neural Networks History	5
2.1.2 The Basic Components of Neural Networks	6
2.1.2.1 Neuron	6
2.1.2.2 Multilayer Perceptron Structure	9
2.1.3 The Characteristics of Neural Networks	11
2.2 Backpropagation Algorithm	12
2.2.1 Complex-valued Backpropagation	13
2.3 Previous Works	15
Chapter 3 System Description	18
3.1 System Model	18
3.1.1 Transmitter	18
3.1.2 Interference	20
3.1.2.1 The CCI Model	20
3.1.2.2 The ACI Model	23
3.1.3 Transmit and Receive Filters	24
3.1.4 AWGN	25
3.2 Equalizer Analysis	25
3.2.1 LMS Decision Feedback Equalizer	29
3.2.2 MLP Decision Feedback Equalizer	32

Chapter 4 Computer Simulation Results of Co-channel Interference	35
4.1 System Model	35
4.2 System Performance Criteria	35
4.2.1 Mean Squared Error	35
4.2.2 Average Probability for Symbol Error	36
4.3 System Performance with a Matched Filter	37
4.3.1 Intersymbol Interference	38
4.3.2 SER Performance without CCI	39
4.3.3 SER Performance without Timing Misalignment of CCI	41
4.3.4 SER Performance with Timing Misalignment of CCI	41
4.4 System Performance with Equalizer	47
4.4.1 System Performance with LMSDFE	47
4.4.1.1 Basic Simulation Parameters	47
4.4.1.2 Convergence Rate	47
4.4.1.3 Equalizer Length	51
4.4.1.4 SER Performance without Timing Misalignment of CCI	53
4.4.1.5 SER Performance with Timing Misalignment of CCI	56
4.4.2 System Performance with MLPDFE	59
4.4.2.1 Basic Simulation Parameters	59
4.4.2.2 BP Parameters and Configuration	60
4.4.2.3 Convergence Rate	66
4.4.2.4 SER Performance without Timing Misalignment of CCI	68
4.4.2.5 SER Performance with Timing Misalignment of CCI	71
4.5 Comparison of SER Performance of Three Receivers	78
4.5.1 SER Performance without Timing Misalignment of CCI	78
4.5.2 SER Performance with Timing Misalignment of CCI	79
 Chapter 5 Computer Simulation results of Adjacent-channel Interference	 82
5.1 System Model	82
5.2 SER Performance with a Matched Filter	82
5.3 SER Performance with Equalizer	84
5.3.1 SER Performance with Two ACIs	84
5.3.2 SER Performance with One ACI	92

Chapter 6 Conclusion and Recommendations	97
6.1 Conclusion	97
6.2 Recommendations for Future Research	98
Appendix A Complex-valued Backpropagation Algorithm	99
A.1 The Generalized Delta Rule	99
A.2 Update of Output Layer Weights	99
A.3 Update of the Hidden Layer Weights	102
Appendix B Complex-Valued LMS Algorithm	104
References	107

List of Figures

Figure 2.1	Adaptive nonlinear combiner, neuron	7
Figure 2.2	Four common threshold functions used in a neuron	8
Figure 2.3	Multilayer perceptron network	10
Figure 3.1	System Block Diagram	19
Figure 3.2	Spectrum of signals after receiver filtering	21
Figure 3.3	Square root raised cosine (SRRC) filter, normalized frequency magnitude response, $\alpha = 0.35$	26
Figure 3.4	Square root raised cosine (SRRC) filter, normalized impulse response, $\alpha = 0.35$	27
Figure 3.5	Raised cosine (RC) filter, normalized impulse response, $\alpha = 0.35$	28
Figure 3.6	LMS Decision-feedback Equalizer	30
Figure 3.7	MLP Decision-feedback Equalizer	33
Figure 4.1	Symbol error rate of a matched filter receiver under no CCI condition	40
Figure 4.2	Symbol error rate of a matched filter as a function of [a] SNR and [b] SIR	42
Figure 4.3	Symbol error rate of a matched filter as a function of the CCI timing misalignment, MIS, at SNR = 18.0db	43
Figure 4.4	Symbol error rate of a matched filter for different MIS as a function of [a] SNR, (SIR = 6dB) and [b] SIR, (SNR = 20dB)	45
Figure 4.5	Symbol error rate of a matched filter as a function of the MIS at [a] SIR = 6dB and [b] SNR = 20dB	46
Figure 4.6	Time evolution of a LMSDFE MSE for [a] SIR = 6 dB SNR= 20dB and [b] SIR = 3 dB SNR = 20dB, MIS = 0.0	50
Figure 4.7	Symbol error rate of a LMSDFE as a function of SNR and DFE configuration at [a] SIR = 6.0 dB [b] SIR = 3.0 dB	52
Figure 4.8	Symbol error rate of a LMSDFE(6, 2) as a function of SNR and LMS step size, μ	54
Figure 4.9	Symbol error rate of a LMSDFE(6, 2)/a matched filter as a function of SNR	55
Figure 4.10	Symbol error rate of a LMSDFE(6, 2) with different MIS as a function of SNR	57
Figure 4.11	Symbol error rate of a LMSDFE(6, 2)/a matched filter as a function of timing misalignment, MIS	58
Figure 4.12	MLPDFE $\overline{mse}$ as a function of momentum parameter, α , at [a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB	62

Figure 4.13 MLPDFE $\overline{mse}$ as a function of Initial BP learning parameter, β_1 , at [a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB	63
Figure 4.14 MLPDFE $\overline{mse}$ as a function of final BP learning parameter, β_1 at [a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB	64
Figure 4.15 MLPDFE $\overline{mse}$ as a function of number of neurons in the first hidden layer, N_1 , at SIR = 3.0 dB SNR = 20.0 dB	65
Figure 4.16 Time evolution of MLPDFE/LMSDFE MSE	67
Figure 4.17 Symbol error rate of an MLPDFE((6, 2), 15, 1) as a function of SNR using [a] on-line training and [b] off-line training	69
Figure 4.18 Symbol error rate of an MLPDFE((6, 2), 15, 1) as a function of SNR, comparing the on-line and off-line training	70
Figure 4.19 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training / a matched filter as a function of [a] SNR [a] and [b] SIR	72
Figure 4.20 Symbol error rate of MLPDFE((6, 2), 15, 1) using off-line training / a matched filter as a function of [a] SNR [a] and [b] SIR	73
Figure 4.21 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training as a function of [a] SNR and [b] MIS at SIR = 6.0 dB	75
Figure 4.22 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training as a function of [a] SNR and [b] MIS at SIR = 4.0 dB	76
Figure 4.23 Symbol error rate of MLPDFE((6, 2), 15, 1) using off-line training as a function of MIS at [a] SIR = 6.0 dB and [b] SNR = 20.0 dB	77
Figure 4.24 SER performance comparisons among three receivers as a function of SNR at [a] SIR = 6.0 dB and [b] SIR = 2.0 dB	79
Figure 4.25 SER performance comparisons among the three receivers as a function of MIS at SIR = 6.0 dB [a] SNR = 16.0 dB and [b] SNR = 18.0 dB	81
Figure 5.1 Symbol error rate of a matched filter for different channel spacing with two ACI	83
Figure 5.2 MLPDFE((9,4),6,1) Symbol error rate for different channel spacing at fixed [a] SIR = 3.64 dB and [b] SIR = 8.2 dB	86
Figure 5.3 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.65$	87
Figure 5.4 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.70$	88
Figure 5.5 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.75$	89
Figure 5.6 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.95$	90
Figure 5.7 MLPDFE/LMSDFE Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$	91

Figure 5.8	MLPDFE((9,4),6,1) Symbol error rate for different channel spacing at fixed [a] SIR = 3.64 dB and [b] SIR = 6.38 dB	93
Figure 5.9	MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$	94
Figure 5.10	MLPDFE/LMSDFE Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$	95

List of Tables

Table 4.1	Basic simulation parameters of LMSDFE	48
Table 4.2	LMSDFE convergence rate simulation parameters	49
Table 4.3	LMSDFE(6, 2) convergence rate at SIR = 6dB and SNR = 20dB	49
Table 4.4	Basic simulation parameters of MLPDFE	59
Table 4.5	MLPDFE BP parameters and configuration simulation parameters	61
Table 5.1	MLPDFE BP parameters and configuration simulation parameters	85

List of Abbreviations and Acronyms

ACI	-Adjacent-channel Interference
AWGN	-Additive White Gaussian Noise
BER	-Bit Error Rate
BP	-Backpropagation Algorithm
BPSK	-Binary Phase Shift Keying
CCI	-Co-channel Interference
CIR	-Channel Impulse Response
CNR	-Carrier-to-Noise Ratio
CW	-Continuous Wave
db	-decibel
DFE	-Decision Feedback Equalizer
DSP	-Digital Signal Processing
FIR	-Finite Impulse Response
GDR	-Generalized Delta Rule
GSM	-Groupe Special Mobile
Hz	-Hertz (cycle/s)
ISI	-Intersymbol Interference
kbps	- 10^3 bits per second
kHz	- 10^3 Hertz
ksps	- 10^3 symbols per second
LE	-Linear Equalizer
LMS	-Least Mean Square
MHz	- 10^6 Hertz
MIS	-Timing Misalignment

MLP	-Multilayer Perceptron
MLPLE	-Multilayer Perceptron-based Linear Equalizer
MLPDFE	-Multilayer Perceptron-based Decision Feedback Equalizer
MLSE	-Maximum Likelihood Sequence Estimate
MMSE	-Minimum Mean Square Error
MSE	-Mean Square Error
PAM	-Pulse Amplitude Modulation
PSK	-Phase Shift Keying
QPSK	-Quadri Phase Shift Keying
RC	-Raised Cosine
RLS	-Recursive Least Square
SER	-Symbole Error Rate
SIR	-Signal-to-Interference Ratio
SNOI	-Signal Not of Interest
SNR	-Signal-to-Noise Ratio
SOI	-Signal of Interest
SRRC	-Square Root Raised Cosine
TDL	-Tapped Delay Line
VLSI	-Very Large Scale Integration

Lists of Mathematical Notation

a_k	-transmitted symbol of the main channel at time $t = kT$
$a_k^{(i)}$	-transmitted symbol of the i^{th} interference channel at time $t = kT$
$a_{kc}^{(i)}$	-transmitted symbol of the i^{th} co-channel at time $t = kT$
$a_{ka}^{(i)}$	-transmitted symbol of the i^{th} adjacent-channel at time $t = kT$
B_r	-the receiver bandwidth where the bandwidth is relative to the symbol rate
B_t	-the transmitter bandwidth where the bandwidth is relative to the symbol rate
D	-delay in the system
D_k	-complex-valued desired output
$f(\cdot)$	-real-valued threshold function
$\dot{f}(\cdot)$	-the derivative of the function $f(\cdot)$
$F(\cdot)$	-complex-valued threshold function
e^2	-sum of the squares of the errors for all output neurons
$E_k^{(M)}$	-complex-valued error of the k^{th} neuron in the output layer
E_s	-the symbol energy
$h[k]$	-the filter impulse response which has unit energy
$h_r(t)$	-the impulse response of the receive filter
$H_r(f)$	-the frequency response of the receive filter
$h_t(t)$	-the impulse response of the transmit filter
$H_t(f)$	-the frequency response of the transmit filter
$H_{rc}(f)$	-the frequency response of the raised cosine filter
$h_{src}(t)$	-the impulse response of the square root raised cosine filter
$H_{src}(f)$	-the frequency response of the square root raised cosine filter

$Im[\]$	-imagery part of a complex value
$mse(nT)$	-the output mean squared error at time $t = nT$
P_s	-probability of symbol error
N_0	-the number of neurons in the input layer
N_b	-the number of feedback taps
N_f	-the number of forward taps
N_l	-the number of neurons in the l th hidden layer
N_M	-the number of neurons in the output layer
$n(t)$	-filtered complex-valued AWGN
$net_j^{(l)}$	-real-valued sum of the inputs to the j th neuron in l th layer
$NET_j^{(l)}$	-complex-valued sum of the inputs to the i th neuron in l th layer
$r(t)$	-received signal at the input of the equalizer
$s(t)$	-the impulse response of the overall filtering
$Re[\]$	-real part of a complex value
$x_i^{(l-1)}$	-real-valued activity level of the i th input neuron
$X_i^{(l-1)}$	-complex-valued activity level of the i th input neuron
X^*	-the complex conjugate of X
$X_k^{(M)}$	-complex-valued actual output of the k th neuron
$\underline{X}$	-complex-valued input vector to the DFE
$w_{ij}^{(l)}$	-real-valued weight coefficient in the l th layer
$W_{ij}^{(l)}$	-complex-valued weight coefficient in the l th layer
$\underline{W}$	-complex-valued weight coefficient vector of the DFE
$\underline{W}_f$	-complex-valued forward weight coefficient vector of the DFE
$\underline{W}_b$	-complex-valued feedback weight coefficient vector of the DFE
T	-symbol period
α	-detected signal of DFE

α	-the BP momentum parameter
β	-the BP learning parameter
β_i	-the initial BP learning parameter
β_f	-the final BP learning parameter
δf	-normalized channel spacing
$\delta[n]$	-Kronecher delta function
$\theta_j^{(l)}$	-the real-valued internal threshold of the j th neuron in the l th layer
$\Theta_j^{(l)}$	-the complex-valued internal threshold of the j th neuron in the l th layer
σ^2	-the variance of data
τ_l	-the timing misalignment of the interference channel with respect to the main channel
μ	-the LMS step size
Δ	-the complex-valued change in weight coefficient
$\Delta_j^{(l)}$	-the complex-valued error signal

Chapter 1

Introduction

Demands for more radio channels have been increasing, due to a rapid increase needs in mobile voice and data communications, and this trend is expected to continue [1–3]. However, allocation of more RF channels is becoming difficult due to the limit in the frequency spectrum availability. The demand for services requires more efficient use of the radio spectrum.

In cellular radio and dual-polarized microwave radio systems, many techniques such as close packing of frequency-division multiplexed signals [4], or smaller distances among frequency re-used systems [5, 6] have been used to improve the spectrum utilization. These approaches often exhibit limitations in performance due to co-channel interference (CCI) and adjacent-channel interference (ACI). CCI is the interference between signals with similar carrier frequencies, whereas ACI is the interference between signals with different carrier frequencies which are close enough to cause spectral overlaps. In addition to ACI and CCI, channel intersymbol interference (ISI) and additive white Gaussian noise (AWGN) are present in radio communications. In general, CCI and ACI limit the system user capacity per unit area, and ISI limits the maximum data rate. Ultimately, all of these impairments can lead to a severe degradation in the performance of the communication system. Thus, the proper communications technique and receiver structure are required to overcome these impairments in order to ensure reliable digital communications over a radio channel.

Adaptive equalizer can be used in these communication systems to obtain reliable data transmission. The optimum receiver structure is the maximum likelihood sequence estimate (MLSE). However, most of the channel equalizers employ a symbol-decision structure such as linear equalizer (LE) and decision feedback equalizer (DFE). Conventional LE and DFE equalizer structures employing adaptive linear algorithms such as least mean square (LMS) and recursive least squares (RLS) have been and are extensively studied and used for various applications [7]. Diversity combining techniques

for these channels have also been and are being intensely researched [8]. Frequency-shift filtering is another technique which can remove co-channel and adjacent-channel interference [9].

The neural network structures and learning algorithm are relatively new techniques with potential advantages in channel equalization and interference reduction. In general, communication systems can be modelled in a linear fashion and thus receivers based on MLSE, LE and DFE can perform sufficiently well. MLSE yields optimum performance in the presence of AWGN, while LE and DFE achieve suboptimum performance [7]. However, ISI can produce nonlinear decision boundaries. Also, in the presence of CCI and ACI, the decision boundaries are generally nonlinear, and this nonlinearity will restrain the performance of the linear receiver. Therefore, for a communication system with an inherent nonlinearity, a receiver employing a neural network with its nonlinear neuron processing can yield better performance than a linear receiver, since the neural network has the ability to form arbitrarily complicated decision boundaries.

1.1 Problem Definition

In a mobile radio communication system, CCI and ACI are two of the major impairments and they limit the number of users per unit. In some applications the problem of ACI is not considered to be a major impairment since it is assumed that this impairment is greatly reduced by filtering and by avoiding the use of adjacent channels by neighboring radio stations. But, these assumptions may result in a loss of valuable spectral capacity. To date, most research of dealing with CCI and ACI in radio communications has been concentrated on techniques such as adaptive equalization and diversity combining to mitigate these channel impairments.

This thesis proposes a multilayer perceptron-based adaptive equalizer (MLP) which has the ability to form complicate nonlinear decision region. The multilayer perceptron-based equalizer is trained by the backpropagation algorithm (BP) to form the decision boundaries. We investigate CCI and ACI separately, and compare the performance of the multilayer perceptron-based equalizer with the that of a DFE equalizer. The timing

misalignment between the desired channel and the co-channel is the major concern in this thesis.

1.2 Thesis Objectives

The objectives of this thesis are as follows:

1. investigate the effect of the MLP configuration and the BP parameters on the performance of a multilayer perceptron-based equalizer in terms of averaged mean square error (MSE) and symbol error rate (SER).
2. evaluate the SER performance of a multilayer perceptron-based equalizer under the condition that the main channel and co-channel are perfectly synchronized and compare its SER performance with that of a matched filter and a conventional DFE.
3. evaluate the SER performance of a multilayer perceptron-based equalizer under the condition that the main channel and co-channel are not synchronized and compare its SER performance with that of a matched filter and a conventional DFE.
4. investigate the minimum channel spacing between main channel and adjacent-channel to achieve certain SER, and compare the SER performance of the multilayer perceptron-based equalizer with that of a DFE.

1.3 Thesis Organization

This thesis is organized as follows:

In Chapter 2, the concept of neural networks is provided: its components, its layer, its characteristics and its learning algorithms. A survey of literatures addressing the same kind of problems and the previous related work is briefly presented.

In Chapter 3, the system model to be used in our simulation is presented. The two types of equalizers, namely Least-Mean-Square decision feedback equalizer (LMSDFE) and multilayer perceptron-based decision feedback equalizer (MLPDFE) are discussed in details.

In Chapter 4, the system performance based on the presence of only co-channel

interference (CCI) is investigated. Comparisons are made among the three different equalizers, namely a matched filter, a LMSDFE and an MLPDFE.

In Chapter 5, the system performance based on the presence of only adjacent-channel interference (ACI) is investigated. Again, comparisons are made among the three different equalizers.

In Chapter 6, concluding remarks are presented along with some recommendations for future research.

Chapter 2

Background Information and Literature Survey

In this chapter, we provide some background information of neural networks: the history of the neural networks, the essential elements (neurons) in neural networks, the multilayer perceptron structure, the characteristics of neural networks and the backpropagation learning algorithm. Finally, a brief survey of previous work related to this thesis is provided.

2.1 An Overview of Neural Networks

The neural network model is an information-processing technology inspired by the studies of the human's brain and nervous system. The model is composed of many non-linear computational elements (neurons) operating in parallel and arranged in patterns reminiscent of biological neural networks. Neurons are connected via weights that are typically adapted during training to improve performance. The studies on neural network models have a long history. The development of detailed mathematical models began more than forty years ago. Only recently, there has been a resurgence of interest in the field caused by new network topologies and learning algorithms, very large scale integration (VLSI) implementation techniques, and the belief that massive parallelism is essential for high performance speech and image recognition. The potential benefits of neural networks extend beyond the high computation rates provided by massive parallelism. Neural networks typically provide a greater degree of robustness or fault tolerance than von Neumann sequential computers because there are many more processing nodes, each with primarily local connections. Damage to a few nodes or links thus need not impair overall performance significantly.

2.1.1 Neural Networks History

Neural networks implementations date back to as early as 1913, when the first neural

networks implementation, using hydraulics, was proposed but never implemented. In 1951, Minsky and Edmonds created the first physical neural networks implementation. In 1957, Frank Rosenblatt [10] began to work on a neural network model called “perceptron”. Built in hardware, it was the oldest neural networks and its concepts are still used today in various forms of applications such as pattern recognition.

More neural networks implementations were announced in 1986 when several major research laboratories developed electronic implementations of neural networks. Among them the feed-forward multilayer perceptron networks implementation was a great achievement. This perceptron model was introduced by Rosenblatt in the late 1950s, but it was not used in the past because effective training algorithms were not available. This has recently changed with the development of a new training algorithm, called backpropagation algorithm [11].

2.1.2 The Basic Components of Neural Networks

2.1.2.1 Neuron

The basic element of neural networks is the neuron (or processing element) which is nonlinear, and typically analog. The neuron sums $N_{l,l}$ weighted inputs and passes the result through a nonlinearity as shown in Figure 2.1. The neuron is characterized by an internal real-valued threshold or offset θ_j and by the type of nonlinearity.

The input signals come from the environment or outputs of other neurons and form an input vector $(x_0^{(l-1)}, \dots, x_i^{(l-1)}, \dots, x_{N_{l-1}-1}^{(l-1)})$, where $x_i^{(l-1)}$ is the real-valued activity level of the i th input neuron in the $(l-1)$ th layer. The collection of weights that abuts the j th neuron in the l th layer, $x_j^{(l)}$, forms a vector $(w_{0j}^{(l)}, \dots, w_{ij}^{(l)}, \dots, w_{(N_{l-1}-1)j}^{(l)})$, where the real-valued weight $w_{ij}^{(l)}$ represents the connection strength from the neuron $x_i^{(l-1)}$ to the neuron $x_j^{(l)}$. The output value $x_j^{(l)}$ is defined as:

$$x_j^{(l)} = f \left(\sum_{i=0}^{N_{l-1}-1} w_{ij}^{(l)} x_i^{(l-1)} + \theta_j^{(l)} \right) \quad (2.1)$$

Figure 2.1 Adaptive nonlinear combiner, neuron

The nonlinearity, also referred to as threshold function, maps a neuron's input domain — to a pre-specified output range. Four common used nonlinearities are the linear, ramp, step, and sigmoid functions. Figure 2.2 shows typical shapes of these four functions. Among them, the sigmoid function is mostly used in neural networks.

The neurons store knowledge or information in the weights $\{w_{ij}\}$ and the weights and the threshold can be fixed or adapted using a number of different algorithms. One neuron can also be considered as a single-layer perceptron and this perceptron has the ability to learn to recognize simple patterns. In [12], Lippmann indicated that the decision regions formed by signal-layer perceptrons are similar to those formed by maximum likelihood Gaussian classifiers which assume inputs are uncorrelated and distributions for different classes differ only in mean values. Neither the single-layer perceptron convergence procedure nor the Gaussian classifier is appropriate when classes cannot be separated by a hyperplane.

$$(a) f(x) = \alpha x$$

$$(b) f(x) = \begin{cases} \gamma & x \geq \gamma \\ x & |x| < \gamma \\ -\gamma & x \leq -\gamma \end{cases}$$

$$(c) f(x) = \begin{cases} \gamma & x > 0 \\ -\delta & \text{otherwise} \end{cases}$$

$$(d) f(x) = (1 - e^{-x}) / (1 + e^{-x})$$

Figure 2.2 Four common threshold functions used in a neuron

2.1.2.2 Multilayer Perceptron Structure

Earlier work on single-layer perceptron neural networks was limited owing to the fact that only linear decision boundaries could be formed [13]. However, later researchers showed that multilayer perceptron neural networks could be used to form much more complex, nonlinear mappings between the input and the output layer. Multilayer perceptron networks are feed-forward networks with one or more layers of neurons between the input and output layers. These additional layers contain hidden units or neurons that are not directly connected to both the input and output neurons. A three-layer network with two layers of hidden units is shown in Figure 2.3. The hidden layers provide the capability to use nonlinear sigmoid's ability to create intricately curved partitions of space. In general, all neurons on a layer are fully inter-connected to the neurons in adjacent layer, but there is no connection within a layer, and normally no connection bridging layers.

When constructing a multi-layer perceptron network, the question of how many hidden layers and how many neurons in each layer are needed is a common question. In general, there should be enough hidden neurons in order to generate a sufficient complex classification mapping. On the other hand, if too many hidden neurons are employed, the convergence of the training process can be very slow. Because a priori knowledge of the optimal classification mapping to be realized is not available, it is very difficult, if not impossible to work out precisely the optimal number of neurons at each hidden layer. In [12], Lippmann did some analysis to prove that a three-layer perceptron neural network could form arbitrarily complex decision regions and could separate the meshed classes. A three-layer perceptron could form regions as complex as those formed using mixture distributions and nearest-neighbor classifiers [14]. His analysis demonstrated that no more than three layers (two hidden layers) were required in perceptron-like feed-forward networks because a three-layer perceptron could generate arbitrarily complex decision regions. He also provided some insight into the problem of selecting the number of neurons to be used in a three-layer perceptron network.

Multilayer perceptrons overcome many of the limitations of single-layer perceptrons,

Figure 2.3 Multilayer perceptron network

but were generally not used in the past because effective training algorithms were not available. This has been changed with the development of a new training algorithm: the backpropagation algorithm which is to be discussed in a separate section in this chapter.

2.1.3 The Characteristics of Neural Networks

- ***Learning***

Neural networks can learn from experience. This means that neural networks can modify their behavior in response to their environment. This factor, more than any other, is responsible for the interest they have received. Being shown a set of inputs (perhaps with desired outputs), they self-adjust to produce consistent responses.

- ***Generalization***

Once trained, a neural network's response can be, to a degree, insensitive to minor variations in its input. This ability to see through noise and distortion to the pattern that lies within is vital to pattern recognition in a real-world environment. Overcoming the literal-mindedness of the conventional computer, it produces a system that can deal with the imperfect world in which we live.

- ***Pattern Recognition***

Pattern recognition tasks require the ability to match large amounts of input information simultaneously and then generate categorical or generalized output. They also provide a reasonable response to noisy or incomplete inputs. Because neural networks learn by being shown examples and the expected results or form their own associations without being prompted and rewarded, they are good at pattern matching types of problems.

- ***Fault Tolerance***

Unlike conventional computers, in neural networks, the information is not stored in a single memory location, rather, it is distributed throughout the system. Therefore, neural networks have robustness capability. In most neural networks, if some neurons are destroyed, or their connections slightly modified, the behavior of the networks as a whole is only slightly degraded. If more neurons are damaged, the behavior of the

networks is degraded slightly further and the performance will suffer, but the system does not come to an abrupt halt.

- ***Limitations***

There is a number of problems that neural networks do not handle well. One of the most important limitations is the learning. Learning, which is based on what the network is taught through its training data, can be tedious. Unfortunately, there are no standards for either a training algorithm or a training data set. Another issue is that a neural network cannot justify its answer as in an expert system.

2.2 Backpropagation Algorithm

Most neural networks algorithms adapt connection weights in time to improve performance based on current results. Adaptation or learning is a major focus of neural networks research. Currently, there are a number of such algorithms in the literature. However, the training algorithms are generally of some variations of the following three types: the Hebbian learning, the Delta rule learning and the competitive learning.

One of the most popular learning algorithms is the backpropagation learning rule [11], which is based on the Delta rule and made possible the training of multi-layer perceptron neural networks in digital signal processing, pattern recognition, and control data compression.

The backpropagation algorithm is a generalization of the LMS algorithm. The LMS algorithm was formulated by Widrow [15] and uses a gradient search technique to minimize a cost function equal to the mean square difference between the desired and the actual network output. The neural network is trained by initially selecting small random weights and internal thresholds and then presenting all training data repeatedly. Weights and thresholds are adjusted after every trial using side information specifying the correct class until weights converge and the cost function is reduced to an acceptable value. An essential component of the algorithm is the iterative method that propagates error terms required to adapt weights back from neurons in the output layer to neurons in lower layers.

2.2.1 Complex-valued Backpropagation

Neural networks have had a wide variety of applications in digital signal processing. In particular the multilayer perceptron networks have been extensively used in nonlinear processing of real-valued coefficients.

Assuming a sigmoid threshold the real-valued output $x_j^{(l)}$ of the neuron j in the l th layer can be expressed as:

$$x_j^{(l)} = f(NE T_j^{(l)}) = \frac{1 - \exp(-NE T_j^{(l)})}{1 + \exp(-NE T_j^{(l)})} \quad (2.2)$$

where

$$NE T_j^{(l)} = \sum_{i=0}^{N_{l-1}-1} w_{ij}^{(l)} x_i^{(l-1)} + \theta_j^{(l)} \quad (2.3)$$

in which

$w_{ij}^{(l)}$ is the real-valued weight coefficients in the l th layer

$\theta_j^{(l)}$ is the real-valued internal threshold of neuron j in the l th layer

$x_i^{(l-1)}$ is the real-valued input of neuron j in the l th layer and

N_{l-1} is the number of inputs $x_i^{(l-1)}$.

The output of the networks can be obtained by repeating the procedure in equation (2.2). In equation (2.2), all the components are real values. The weight coefficients can be adjusted using the real-valued backpropagation algorithm [11].

The real-valued backpropagation above can be used to solve a real-valued signaling system. However, in many applications, a complex-valued mathematical representation of signaling is required. For instance, the QPSK signal has two components, namely, amplitude and phase. In [16], Benvenuto proposed a complex-value backpropagation algorithm to include complex signal and coefficients. For a complex-valued signal, the summation of inputs to the neuron, $NE T_j^{(l)}$, can be obtain by equation (2.3) while all the

inputs and coefficients are complex values. The sigmoid function for the complex-value [16], can be shown as:

$$X_j^{(l)} = F(NE T_j^{(l)}) = f\left(\operatorname{Re}\left[NE T_j^{(l)}\right]\right) + j f\left(\operatorname{Im}\left[NE T_j^{(l)}\right]\right) \quad (2.4)$$

where

$\operatorname{Re}\left[NE T_j^{(l)}\right]$ is the real part of the complex-valued $NE T_j^{(l)}$

$\operatorname{Im}\left[NE T_j^{(l)}\right]$ is the imagery part of the $NE T_j^{(l)}$ and

$f(\cdot)$ is the conventional real-valued sigmoid function defined in equation (2.2).

The complex BP algorithms (see appendix A for a derivation) can be summarized as follows:

1. Initialize weight coefficients and internal threshold coefficients

Set all weights and thresholds to small random values

2. Present input data and desired output data

Present input vector $\left[X_0^{(0)}, \dots, X_i^{(0)}, \dots, X_{N_0-1}^{(0)}\right]$ and the desired response vector $\left[D_0, \dots, D_k, \dots, D_{N_M-1}\right]$ (here N_0 is the number of neurons in the input layer and N_M is the number of neurons in the output layer (M th layer))

3. Calculate the actual networks output

Use the sigmoid function defined in equation (2.4) to calculate the output vector $\left[X_0^{(l)}, \dots, X_j^{(l)}, \dots, X_{N_l-1}^{(l)}\right]$ in the l th layer. The output of neurons in the l th layer is the input to the neurons in the $(l+1)$ th layer and it multiplies with the weights in the $(l+1)$ th layer to produce the output of neurons in the $l+1$ layer. Eventually, the output vector $\left[X_0^{(M)}, \dots, X_k^{(M)}, \dots, X_{N_M-1}^{(M)}\right]$ of the neurons in the output layer will be obtained

4. Adapt weights and thresholds

Use a recursive algorithm starting at the output neurons and working your way back to the first hidden layer. Adjust weights by: $l = M, \dots, 0$ and $j = 0, \dots, N_l - 1$

$$\begin{aligned} W_{ij}^{(l)}((n+1)T) = & W_{ij}^{(l)}(nT) + \beta \Delta_j^{(l)} X_i^{*(l-1)} \\ & + \alpha \left[W_{ij}^{(l)}(nT) - W_{ij}^{(l)}((n-1)T) \right] \end{aligned} \quad (2.5)$$

where

* is the complex conjugate

β is the learning-rate parameter and is analogous to the LMS step size, μ

α is the momentum parameter and its purpose is to speed up convergence by reducing the time variations in the weights during adaptation and

the error signal $\Delta_j^{(l)}$ is:

$$\Delta_j^{(l)} = \begin{cases} \left(\operatorname{Re}[D_j] - \operatorname{Re}[X_j^{(l)}] \right) f' \left(\operatorname{Re}[NET_j^{(l)}] \right) + j \left(\operatorname{Im}[D_j] - \operatorname{Im}[X_j^{(l)}] \right) f' \left(\operatorname{Im}[NET_j^{(l)}] \right), & \text{for } l = M \\ \sum_{k=0}^{N_l-1} W_{kj}^{*(l)} \left(\operatorname{Re}[E_j^{(l)}] f' \left(\operatorname{Re}[NET_j^{(l-1)}] \right) + j \operatorname{Im}[E_j^{(l)}] f' \left(\operatorname{Im}[NET_j^{(l-1)}] \right) \right), & \text{for } l \neq M \end{cases} \quad (2.6)$$

where ' is the derivative of sigmoid function $f(\cdot)$ in equation (2.2) and

$E_j^{(l)}$ is the error for the j th neuron in the l th layer (see equation (A.15)).

5. Repeat by going to step 2

2.3 Previous Works

In the past decade, neural networks have been applied to many diverse fields in signal processing. A neural network which has the ability to learn sophisticated non-linear relationships provides processing capability beyond the limits set by linear adaptive filters.

The multilayer perceptron network using backpropagation learning algorithm is a versatile tool that is readily applied to a number of diverse problems. To a large extent, its versatility is due to the general nature of the network learning process. The generality offered by this common process allows arrangement and connectivity of individual units within the network that can vary dramatically. Similarly, due to the variety of the network structures that can be created and trained successfully using the backpropagation algorithm, this learning technique can be applied to many different kinds of problems.

The multilayer perceptron network using backpropagation algorithm is a powerful adaptive method for filtering out noise or identifying underlying signals. There are many

papers studying the capacity of this kind of network. Klimasanakas [17] discussed the use of a three-layer feed-forward perceptron that employed the backpropagation learning rule for noise filtering. His simulation results showed that a multilayer perceptron network technique was a powerful method for noise filtering and it offered an interesting and potentially powerful approach.

Gilbson et.al. [18] have previously investigated a non-linear equalizer structure based on a multilayer perceptron network and demonstrated its superior performance over the conventional LMSLE. Siu et.al. [19] extended this work to the case of decision-feedback equalization. Both studies claimed that a MLP-based equalizer could outperform a conventional equalizer of similar ISI span in the presence of AWGN. These previous investigations looked at the case of real-valued and T -spaced systems with fixed learning parameter values. Also none of these works took the effect of CCI and ACI into account.

In his course project, Lo [20] investigated the performance of a three-layer perceptron equalizer employing a complex-valued backpropagation learning algorithm when the CCI was present in the system. In his study, $\frac{T}{2}$ -spaced system was used and variable learning parameters were employed to improve the equalizer convergence rate. His simulation results indicated that MLPDFE could essentially match the bit error rate (BER) performance of a conventional LMSDFE under all noise and interference conditions.

There is, however, no literature which studied the system performance under ACI environment using neural network techniques, even though this problem has been addressed in many papers [21, 22]. Niger and Vandamme studied the performance of equalization techniques for digital radio system in the presence of Continuous Wave (CW) interference and ACI [22]. Their work indicated that $\frac{T}{2}$ -spaced and decision feedback equalizers could improve ACI margin significantly, and their results were given in terms of degradation of carrier-to-noise ratio (CNR) due to ACI for fixed channel spacing. The potential difference on system performance between the case when ACI power was changed due to different channel spacing or due to its level for fixed channel spacing had not been investigated.

Uesugi et al. used computer simulation to show that the signal distortion caused by frequency selective fading and ACI could be removed by fractionally spaced DFE [23]. The channel spacing was assumed to be fixed and the effect of ACI on BER was presented for the case of groupe special mobile (GSM). The effect of channel spacing has not been considered in this work.

S. Golestaneh [24] investigated the degradation of a given signal called signal of interest (SOI) in presence of these adjacent channels signals called signal not of interest (SNOI) as a function of channel spacing and receiver structure. A $\frac{T}{2}$ -spaced DFE was used to reduce channel spacing between adjacent channels under the condition that it would not increase channel distortion and the ISI. His simulation results indicated that for the same signal-to-interference ratio (SIR), a large channel spacing always led to a better SER. Also for the same SIR, the equalizer had better SER performance dealing with one ACI than with two ACI.

Chapter 3

System Description

In this chapter, a block diagram of a general radio system is provided followed by a detailed description of each component. Following this, the structures and the specific adaptation algorithms of two different types of equalizers, namely LMSDFE and MLPDFE, are discussed.

3.1 System Model

Figure 3.1 illustrates a general block diagram of the major components of the radio channel system we consider: the transmitted data $\{a_k\}$, the transmit filter, the channel, the interference $\{a_k^{(i)}\}$ ($i = \dots, -2, -1, 0, 1, 2, \dots$), the additive white Gaussian noise (AWGN), the receive filter and the equalizer.

In this system, we assume that the communication is digital. The data is transmitted using a linear modulation scheme, such as quadrature phase-shift keying (QPSK). Since this model includes the complex baseband representation, the interference can include both co-channel interference (CCI) and adjacent-channel interference (ACI). In the frequency domain the co-channel interference would be centered at 0 Hz, the adjacent-channel interference would be centered at some set of non-zero carrier frequencies so that it would exhibit only vestigial components in the baseband frequencies.

3.1.1 Transmitter

The transmitted data, a_k , shown in Figure 3.1, are transmitted at a bit rate of 48 *kbits/s* over a 30 *kHz* channel. The transmitted data have the properties that they are complex-valued, have zero mean, have unit variance, and are mutually uncorrelated in time:

$$E[a_k] = 0 \quad (3.1)$$

$$E[a_k a_n^*] = \sigma_a^2 \delta[k - n] \quad (3.2)$$

Figure 3.1 System Block Diagram

The variance of the data is noted by:

$$\sigma_a^2 = 1 \quad (3.3)$$

the *Kronecker delta function* is denoted by:

$$\delta[n] = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \quad (3.4)$$

the symbol $E[\]$ denotes expectation of the quantity over the ensemble of a_k , and the symbol $*$ denotes complex conjugate transpose.

3.1.2 Interference

As we mentioned earlier, the interference includes both ACI and CCI signals. The power spectral density of the signal which carries the data of interest, the ACI signals, and the CCI signals after the receive filter are shown in Figure 3.2. The signals which carry the data of interest and the CCI signals are centered at 0 Hz and the ACI signals are centered at non-zero carrier frequencies. The CCI is not separately shown in Figure 3.2 since it is assumed to occupy the same frequency band as the signal which carries the data of interest.

Figure 3.2 introduces three parameters. The transmitter bandwidth, B_t , receiver bandwidth, B_r , and carrier spacing, δf , all of which are measured relative to the symbol rate, $1/T$. Therefore the absolute bandwidths and frequencies may be obtained from the relative bandwidths and frequencies by multiplication with $1/T$. In our simulation system model, the receiver bandwidth is the same as the transmitter bandwidth. This will be discussed in detail in a later section.

3.1.2.1 The CCI Model

The co-channel data, $a_{kc}^{(i)}$, shown in Figure 3.1, are assumed to be the same form (QPSK modulated) and the same bit rate as that of the main channel's data but independent of the main channel's data and of each other. They have the following properties:

$$E[a_{kc}^{(i)}] = 0 \quad (3.5)$$

Figure 3.2 Spectrum of signals after receiver filtering

$$E \left[a_{kc}^{(i)} a_{nc}^{*(i)} \right] = \sigma_a^2 \delta(k - n) \quad (3.6)$$

$$E \left[a_k a_{nc}^{(i)} \right] = 0 \quad (3.7)$$

$$E \left[a_{kc}^{(i)} a_{kc}^{(l)} \right] = 0 \quad (3.8)$$

Superscript (i) refers to the i th co-channel. In our system model, however, only one co-channel interference has been employed (*i.e.*, $i = 1$). The co-channel signal at the input to the equalizer can be written as:

$$x^{(1)}(t) = \sum_{k=-\infty}^{\infty} a_{kc}^{(1)} s^{(1)}(t - kT - \tau_1) \quad (3.9)$$

where $s^{(1)}(t)$ is the impulse response of the overall filtering and can be written as:

$$s^{(1)}(t) = A_1 h_t(t) \otimes h_r(t) \quad (3.10)$$

in which

A_1 corresponds to the CCI attenuation

$h_t(t)$ is the impulse response of the transmit filter

$h_r(t)$ is the impulse response of the receive filter

$\otimes$ denotes the convolution and

τ_1 represents the timing misalignment of the co-channel with respect to the main channel and is assumed to be uniformly distributed on $[0, T]$.

At the input to the equalizer, the received signal can be written as:

$$r(t) = \sum_{k=-\infty}^{\infty} a_k s(t - kT) + \sum_{p=-\infty}^{\infty} a_{kc}^{(1)} s^{(1)}(t - pT - \tau_1) + n(t) \quad (3.11)$$

where

a_k is the transmitted signal on the main channel

$a_{kc}^{(1)}$ is the transmitted signal on the co-channel

$s(t)$ is the impulse response of the overall main channel and

$n(t)$ is the AWGN, after filtering by the receive filter.

For the CCI, the signal to interference ratio (SIR) at the input to the equalizer is defined as:

$$SIR = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_c^2} \right) \quad (3.12)$$

where

σ_s^2 is the average desired signal power and

σ_c^2 is the average co-channel interference power.

3.1.2.2 The ACI Model

The data sequence of the adjacent-channel $a_{ka}^{(i)}$ are assumed to be of the same form (i.e. QPSK modulated) and the same bit rate as that of the main channel's data but independent of the main channel's data and of each other. Each of the adjacent-channel interference signals at the input of the equalizer can be written in its low pass equivalent form as:

$$x^{(i)}(t) = \sum_{k=-\infty}^{\infty} a_{ka}^{(i)} s^{(i)}(t - kT - \tau_i) \quad (3.13)$$

where $s^{(i)}(t)$ is the impulse response of the overall filtering and can be written as:

$$s^{(i)}(t) = A_i h_t(t) e^{j(2\pi\delta f t + \varphi_i)} \otimes h_r(t) \quad (3.14)$$

in which

A_i corresponds to ACI level

δf is the channel spacing (normalized with respect to the symbol duration T)

$h_t(t)$ is the impulse response of the transmit filter

$h_r(t)$ is the impulse response of the receive filter

$\otimes$ denotes the convolution and

τ_i represents the timing misalignment of the adjacent-channel with respect to the main channel and is assumed to be uniformly distributed on $[0, T]$.

3.1.3 Transmit and Receive Filters

Square root raised cosine filters are used as the transmit and receive filters in our simulation model. It is assumed that the overall filtering is equally split between the transmitter and the receiver. Therefore, both the transmit and receive filters are identical:

$$|H_t(f)| = |H_r(f)| \quad (3.15)$$

For square root raised cosine filter, it has the frequency response of the form:

$$H_{src}(f) = \begin{cases} T, & 0 \leq |f| \leq \frac{(1-\alpha)}{2T} \\ T\sqrt{\frac{1}{2}\{1 - \sin[\pi\frac{T}{\alpha}(|f| - \frac{1}{2T})]\}}, & \frac{(1-\alpha)}{2T} \leq |f| \leq \frac{(1+\alpha)}{2T} \\ 0, & |f| > \frac{(1+\alpha)}{2T} \end{cases} \quad (3.16)$$

Alternately, the square root raised cosine filter can be expressed in terms of its impulse response which can be shown to be:

$$h_{src}(t) = \begin{cases} 1 - \alpha + 4\frac{\alpha}{\pi}, & t = 0 \\ \frac{\alpha}{\sqrt{2}}\{(1 + \frac{2}{\pi})\sin(\frac{\pi}{4\alpha}) + (1 - \frac{2}{\pi})\cos(\frac{\pi}{4\alpha})\}, & t = \pm\frac{T}{4\alpha} \\ \frac{\sin[\pi(1-\alpha)\frac{t}{T}] + 4\alpha\frac{t}{T}\cos[\pi(1+\alpha)\frac{t}{T}]}{\pi\frac{t}{T}[1-(4\alpha\frac{t}{T})^2]}, & \text{for all other } t \end{cases} \quad (3.17)$$

where α is the roll-off factor which is between 0 and 1, and T is the symbol period. The overall transfer function of two square root raised cosine filters will have the raised cosine characteristics:

$$H_{rc}(f) = \begin{cases} 1, & |f| \leq \frac{1}{2T}(1 - \alpha) \\ \frac{1}{2}\{1 - \sin[\frac{T}{2\alpha}(2\pi|f| - \frac{\pi}{T})]\}, & \frac{1}{2T}(1 - \alpha) < |f| \leq \frac{1}{2T}(1 + \alpha) \\ 0, & |f| > \frac{1}{2T}(1 + \alpha) \end{cases} \quad (3.18)$$

which meets the Nyquist criteria for zero intersymbol interference [25]. Figure 3.3, Figure 3.4, and Figure 3.5 show the frequency response of a square root raised cosine

filter, the time impulse response of a square root raised cosine filter (the value in equation (3.16) is divided by 2.8), and the time impulse response of a raised cosine filter, respectively.

3.1.4 AWGN

A complex-valued additive white Gaussian noise, $n(t)$, is used to model the system noise. The AWGN has a two-sided power spectral density N_0 and satisfies:

$$E[n(t)] = 0 \quad (3.19)$$

$$E[n(t_1)n^*(t_2)] = N_0\delta(t_1 - t_2) \quad (3.20)$$

The signal-to-noise ratio (SNR) per transmitted symbol after receiver is defined as:

$$\begin{aligned} SNR &= 10\log_{10}\left(\frac{E_s}{N_0}\right) \\ &= 10\log_{10}\left(\frac{\sigma_s^2 \sum_k h^2[k]}{\sigma_n^2}\right) \\ &= 10\log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right) \end{aligned} \quad (3.21)$$

where

E_s is the energy per symbol

σ_s^2 is the variance (power) of the desired signal which is set to unity in the simulation and

$h[k]$ is the filter impulse response which has unit energy.

3.2 Equalizer Analysis

The block diagram in Figure 3.1 does not show any details of the equalizer. In this thesis, two types of equalizers are considered. Figure 3.6 and Figure 3.7 show details

**Figure 3.3 Square root raised cosine (SRRC) filter,
normalized frequency magnitude response, $\alpha = 0.35$**

Figure 3.4 Square root raised cosine (SRRC) filter, normalized impulse response, $\alpha = 0.35$

Figure 3.5 Raised cosine (RC) filter, normalized impulse response, $\alpha = 0.35$

of these two types of equalizers: the decision feedback equalizer using the least mean square algorithm (LMSDFE), and the multilayer perceptron-based decision feedback equalizer (MLPDFE), respectively.

At the input to the equalizer, the received signal is sampled at a rate of $2/T$. So the tap spacing of the forward filter is $T/2$. The advantages of using a fractionally spaced equalizer over a baud rate (T -spaced) equalizer have been studied by many researchers for various applications [22, 23, 25]. Since the sampling rate of a $\frac{T}{2}$ -spaced equalizer is twice as that of a T -spaced equalizer, the periodicity of the spectrum of a $\frac{T}{2}$ -spaced equalizer is $T/2$ while that of a T -spaced equalizer is $1/T$ [22]. This is specially important for timing phase errors.

3.2.1 LMS Decision Feedback Equalizer

The decision-feedback equalizer (DFE) is a nonlinear equalizer which can perform much better than the linear equalizer (LE). The LE equalizer is the simplest equalizer structure and it essentially attempts to “invert” the folded spectrum of the overall channel impulse response (CIR) so as to satisfy the Nyquist criterion. It is well known that the LE suffers from unacceptable noise enhancement when the channel experiences severe amplitude distortion.

The DFE is a technique used in the communication of discrete-time and discrete-level data over a noisy channel with a non-ideal impulse response. It consists of equalizing the channel to produce a sampled causal response, and substituting that part of the intersymbol interference at the sample time that is due to past data decisions. The advantage of this nonlinear receiver is that a portion of the noise enhancement inherent in linear equalization is avoided by using past decisions to subtract out a portion of the intersymbol to propagate because they result in residual intersymbol interference and a reduced margin against noise at future decisions. The algorithm used to adjust the weights of DFE in our model is least mean square algorithm (LMS).

The DFE is comprised of a forward transversal filter, $\underline{W}_f(nT)$, (which is essentially an LE) and a feedback transversal filter, $\underline{W}_b(nT)$ (see Figure 3.6). The forward filter,

Figure 3.6 LMS Decision-feedback Equalizer

$\underline{W}_f(nT)$, is implemented as a $\frac{T}{2}$ -spaced tapped delay line (TDL) with N_f tap coefficients. On the other hand, since a detected signal, $\hat{\underline{\alpha}}(nT)$, is generated at every symbol period, the feedback filter, $\underline{W}_b(nT)$, is implemented as a T -spaced TDL with N_b tap coefficients. With both a forward and a feedback filter, a DFE has more degrees of freedom to process the ISI in the received signal. Specifically, the forward filter removes the precursor ISI so that zero crossings occur at all samples of the received signal prior to the main lobe at the current sample instant. In essence, the forward filter accomplishes this by shifting most of the received signal energy in the anti-causal part of the received signal response (i.e., precursor) to the causal part of the received signal response (i.e., postcursor). Consequently, the postcursor ISI at the samples after the main lobe of the received signal response can be eliminated by subtracting the feedback-filtered symbols from the output of the forward filter.

The output of the equalizer can be written as:

$$Y(nT) = \underline{x}'(nT)\underline{W}(nT) \quad (3.22)$$

where ' is the vector transpose operator and the DFE weight vector is comprised of both forward and feedback tap coefficients:

$$\underline{W}(nT) = \begin{bmatrix} \underline{W}_f(nT) \\ \underline{W}_b(nT) \end{bmatrix} \quad (3.23)$$

while the overall input signal to the DFE is comprised of both received signals, $\underline{r}(nT)$, and detected feedback signals, $\hat{\underline{\alpha}}(nT)$:

$$\underline{x}(nT) = \begin{bmatrix} \underline{r}(nT) \\ \hat{\underline{\alpha}}(nT) \end{bmatrix} \begin{bmatrix} r(nT) \\ r(nT - \frac{T}{2}) \\ \vdots \\ r(nT - (N_f - 1)\frac{T}{2}) \\ \hat{\alpha}((n - D - 1)T) \\ \hat{\alpha}((n - D - 2)T) \\ \vdots \\ \hat{\alpha}((n - D - N_b)T) \end{bmatrix} \quad (3.24)$$

The weight coefficients, $\underline{W}(nT)$, can be updated as (see appendix B):

$$\underline{W}((n + 1)T) = \underline{W}(nT) + \mu(nT)\varepsilon(nT)\underline{x}^*(nT) \quad (3.25)$$

where

$\mu(nT)$ is the LMS adaptation step size

$*$ is the complex conjugate and

$\epsilon(nT)$, the adaptation error signal, is the difference between the detected data, $\hat{a}((n-D)T)$ (after training, D is the system delay) or transmitted data, a_{n-D} (during training) and the output of the equalizer $Y(nT)$.

The LMS step size may be fixed or variable during adaptation. For instance, to improve the convergence performance, the step size during training could be gradually reduced to a suitably small value which is then fixed for the duration of the data transmit mode. We will discuss how to choose suitable LMS step size in detail in the following chapter.

3.2.2 MLP Decision Feedback Equalizer

The MLP-based decision feedback equalizer consists of a two-layer perceptron whose input signal samples are the same as that for the LMSDFE. That is, the input signal to MLPDFE consists of the $N_f \frac{T}{2}$ -spaced forward TDL and the $N_b T$ -spaced backward TDL, respectively (see Figure 3.7). The input data layer does not count as a perceptron layer since it is simply a data buffer without any weight coefficients or nonlinear neuron functions. The input layer has $N_f + N_b$ nodes while the first hidden layer has N_I nodes. The MLP equalizer generates at symbol rate a signal output which is the input to the decision device. This equalizer is denoted as MLPDFE($(N_f, N_b), N_I, 1$).

The output of the first hidden layer can be written as:

$$X_j^{(1)}(nT) = F(NE T_j^{(1)}) \quad (3.26)$$

where

$$\begin{aligned} NE T_j^{(1)}(nT) = & \sum_{i=0}^{N_f + N_b - 1} W_{ij}^{(1)}(nT) r\left(nT - \frac{iT}{2}\right) \\ & + \sum_{i=N_f}^{N_f + N_b - 1} W_{ij}^{(1)}(nT) \hat{a}(nT - DT - (i - N_f)T) + \Theta_j^{(1)}(nT) \end{aligned} \quad (3.27)$$

Figure 3.7 MLP Decision-feedback Equalizer

in which

$W_{ij}^{(1)}(nT)$ is the complex-valued weight coefficient in the first hidden layer

$\Theta_j^{(1)}(nT)$ is the threshold of the j neuron in the first hidden layer

$r(nT)$ is the received signal

$F(\cdot)$ is the complex-valued nonlinear function defined in equation (2.4)

The output of the MLPDFE can be written as:

$$X_k^{(M)}(nT) = F\left(\sum_{j=0}^{N_1-1} W_{jk}^{(M)}(nT) X_j^{(1)}(nT) + \Theta_k^{(M)}(nT)\right) \quad (3.28)$$

where

$W_{jk}^{(M)}(nT)$ is the weight coefficient in the output layer and

$\Theta_k^{(M)}(nT)$ is the threshold of the k neuron in the output layer.

The MLP weight coefficients are adjusted by the complex-valued backpropagation algorithm (see appendix A for the derivation). In our simulation model, since a QPSK modulator is used, all the signals should be considered as complex values. The complex-valued weight coefficients adaptation form is shown in equation (2.5).

Chapter 4

Computer Simulation Results of Co-channel Interference

This chapter presents the simulation results of three receivers, namely a matched filter receiver, a LMSDFE receiver and an MLPDFE receiver, in a white gaussian noise channel corrupted by CCI. The system performance is mainly evaluated in terms of symbol error rate (SER) under two scenarios: one is the system without CCI timing misalignment; the other is the system with CCI timing misalignment. The system model and system performance criteria are first introduced, followed by a presentation and discussion of the simulation results. Finally, the comparison of the SER performance of these three receivers have been made.

4.1 System Model

In chapter 3, the general block diagram of the simulation system model shown in Figure 3.1 has been described. In this chapter, the channel is modelled as an AWGN channel and the effects of ISI, AWGN and CCI on the system performance will be investigated assuming that the ACI is negligible after receiver filtering. For the simplicity of simulations, only a single-path AWGN channel and one CCI is considered. The effect of ACI on system performance will be discussed in the following chapter.

4.2 System Performance Criteria

4.2.1 Mean Squared Error

A common and obvious measure of performance for a digital communications system is the symbol error rate (SER) for a given signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) at the receiver. This SER is measured at the output of the receiver decision device. In addition, the mean squared error (MSE) at the output of the equalizer, but before the decision device, is also a good measure of the

convergence rate of different receivers. Since the reference locations of the MSE and SER measurements are separated by a nonlinear decision device, the relationship between the MSE and SER is highly nonlinear. Despite their nontrivial relationship, the MSE and SER are sufficiently correlated that the MSE can also provide a somewhat indirect, but meaningful, measure of system performance.

In our investigation of the communication system, the equalizer output mean squared error, averaged over an ensemble of R trials, is computed for each of the N samples:

$$mse(nT) = \frac{1}{R} \sum_{k=0}^{R-1} |a_{d,r}((n - D)T) - y_r(nT)|^2 \quad (4.1)$$

where

$a_{d,r}$ is the n th transmitted data in the r th simulation trial

$y_r(nT)$ is the n th actual equalizer output in the r th simulation trial

D is the system delay

From this ensemble statistic, a useful measure of equalizer performance is computed as the average MSE. Specifically, the average MSE is the time average of the MSE over all output samples in a given interval:

$$\overline{mse} = \frac{1}{N} \sum_{n=0}^{N-1} mse(nT) \quad (4.2)$$

The average MSE performance criteria is used to select the approximately optimum simulation parameters. Ideally, these parameter values should optimize the system SER. Since the simulations are very time consuming, the average MSE performance criteria is used instead to quickly obtain some suitable MLP equalizer parameters for subsequent SER simulation studies. Thus, the minimization of the average MSE is a valid criteria for the design of an MLP equalizer.

4.2.2 Average Probability for Symbol Error

The average SER, as a function of the SNR and SIR at the input to an equalizer, is the most common measure of performance for a digital communication system. For the

case of BPSK signaling, the average BER, $\overline{P}_e$, can be written as:

$$\overline{P}_e = \frac{1}{R} \sum_{r=0}^{R-1} e_r \quad (4.3)$$

with

$$e_r = \begin{cases} 0, & \text{if } a_n = \hat{a}_n \text{ (no bit errors)} \\ 1, & \text{otherwise (at least one bit error)} \end{cases} \quad (4.4)$$

where a_n and $\hat{a}_n$ are BPSK data bits at the source and output of the decision device, respectively.

For the case of QPSK signaling where two information bits are carried by one transmitted symbol, the average SER, $\overline{P}_s$, can be approximately expressed as:

$$\overline{P}_s = 1 - (1 - \overline{P}_e)^2 = \overline{P}_e(2 - \overline{P}_e) \approx 2\overline{P}_e \quad (4.5)$$

where, the average SER for QPSK can be measured as:

$$\overline{P}_s = \frac{1}{R} \sum_{r=0}^{R-1} e_r \quad (4.6)$$

4.3 System Performance with a Matched Filter

We start to discuss the results of our simulations first by looking at the system performance when only a matched filter is used as the receiver. In a static Gaussian channel, matched ISI-free filters are used for optimum SER performance [26]. In a matched ISI-free system, ISI is introduced by the transmitter filter; the receiver filter removes the ISI so that the cascade system is ISI-free. In our simulation system, a square root raised cosine filter with 35% rolloff factor is used as the transmit filter and an additive white Gaussian noise channel is employed. Since a square root raised cosine filter with 35% rolloff factor used as a receive filter satisfies the Nyquist criterion, this receive filter is a matched filter.

4.3.1 Intersymbol Interference

Intersymbol interference arises in all pulse-modulation system and it is characterized by smearing of the transmitted pulses such that multiple adjacent symbols contribute to the received signal at each sampling instant [7, 25]. The phenomenon of ISI can be illustrated with (complex) baseband pulse-amplitude modulated (PAM) signaling. The overall channel impulse response $s(t)$ includes the effects of the transmit filter, the modulator, the transmission medium, and the demodulator. The received signal, $r(t)$, is the superposition of the impulse responses of the channel to each transmitted symbol and received-filtered noise, $n(t)$:

$$r(t) = \sum_k a_k s(t - kT) + n(t) \quad (4.7)$$

At the sampling instant, $t = mT + \tau$, the sampled received signal becomes:

$$r(mT + \tau) = a_m s(\tau) + \sum_{k=m} a_k s((m - k)T + \tau) + n(mT + \tau) \quad (4.8)$$

where

the first ($k = m$) term is the desired received signal

the second term is the summation of undesired ISI components

the third term is the sampled received-filtered noise and

τ is the channel propagation delay

From the above equation, it is evident that reliable detection of the received symbol, $a_m s(\tau)$, is possible only if the additive noise and ISI are relatively small at the sampling instants so that the “eye” is sufficiently open. The condition of no ISI requires the overall $s(t)$ to have zero crossings at all sampling instants other than at the desired sampling instant:

$$s((m - k)T + \tau) = \begin{cases} 1 & m = k - \frac{\tau}{T} \\ 0 & \text{otherwise} \end{cases} \quad (4.9)$$

This zero-ISI condition corresponds to the Nyquist criterion in the frequency domain:

$$\sum S\left(f - \frac{k}{T}\right) = \text{constant} \quad |f| \leq \frac{1}{2T} \quad (4.10)$$

where $W = \frac{1}{2T}$ is commonly referred to as the Nyquist or minimum bandwidth.

An ideal lowpass filter with minimum bandwidth satisfies the Nyquist criteria while achieving optimum bandwidth efficiency. Lowpass filters with excess bandwidth and odd symmetry about $\frac{1}{2T}$ Hz also satisfy the Nyquist criterion. The widely used bandlimited filter in practice is the raised-cosine family with cosine rolloff around $\frac{1}{2T}$ Hz. In this thesis, a square root raised cosine filter is used as the transmit and receive filter.

4.3.2 SER Performance without CCI

As a benchmark for the simulation results, the QPSK modulation SER performance of a matched filter under no CCI is first investigated.

The theoretical probability of symbol error P_s of the QPSK signalling over a static AWGN channel with a single receiver is [25]:

$$P_s = \text{erfc}\left(\sqrt{\frac{E_s}{2N_0}}\right) - \frac{1}{4}\text{erfc}^2\left(\sqrt{\frac{E_s}{2N_0}}\right) \quad (4.11)$$

where

E_s is the symbol energy

$\frac{N_0}{2}$ is the power spectral density of the AWGN and

$\text{erfc}(\cdot)$ is the complementary error function defined as:

$$\text{erfc}(t) = \frac{2}{\sqrt{\pi}} \int_t^{\infty} \exp(-x^2) dx \quad (4.12)$$

Since the simulation results are presented as SER versus SNR curves, the probability of symbol error can be expressed in terms of SNR (defined in equation (3.21)) as:

$$P_s = \text{erfc}\left(\sqrt{\frac{10^{\frac{\text{SNR}}{10}}}{2}}\right) - \frac{1}{4}\text{erfc}^2\left(\sqrt{\frac{10^{\frac{\text{SNR}}{10}}}{2}}\right) \quad (4.13)$$

Figure 4.1 is the simulation result of a matched filter while only AWGN appears. The result is in close agreement with the theoretical result obtained from the above

Figure 4.1 Symbol error rate of a matched filter receiver under no CCI condition

equation (4.13) and it also verifies the correctness of the simulation software. Both the simulated and the theoretical SER curves display the characteristics “waterfall” drop with increasing SNR values. This curve is the optimum result for an AWGN signal-path channel.

4.3.3 SER Performance without Timing Misalignment of CCI

In this section, the SER performance of a matched filter is simulated for a channel corrupted by both AWGN and co-channel interference. Only one co-channel interference is considered here. As we discussed in chapter 2, the system may have CCI timing misalignment, τ_I , with respect to the main channel signal. In this section, we assume that the co-channel signal and the main channel signal are perfectly synchronized, i.e., $\tau_I = 0.0$.

The effect of co-channel interference on SER performance of a matched filter is shown in Figure 4.2 . Figure 4.2 [a] is the SER performance as a function of SNR and Figure 4.2 [b] as a function of SIR. It can be easily viewed from this figure that the increasing of co-channel interference degrades the simulated SER performance. Figure 4.2 [b] clearly shows the effect.

4.3.4 SER Performance with Timing Misalignment of CCI

In the above section, all the simulations are done under the assumption that the main channel and co-channel are perfectly synchronized. However, in reality, this assumption is not always valid, since a delay between main and co-channel does exist. In this section, assuming that the delay is uniformly distributed in the range of $[0, T]$, we investigate the effect of timing misalignment of CCI on the SER performance of a matched filter.

The normalized timing misalignment of CCI by symbol period, T , is presented in terms of MIS . When we say that $MIS = 0.5$, it means that we assume that the timing misalignment of CCI in the communication system is fixed at this value, therefore all the received signals have the same MIS . Figure 4.3 shows the SER performance of a matched filter as a function of the timing misalignment, MIS . Form this figure, we

Figure 4.2 Symbol error rate of a matched filter as a function of [a] SNR and [b] SIR

Figure 4.3 Symbol error rate of a matched filter as a function of the CCI timing misalignment, MIS, at SNR = 18.0db

observe that the plot is almost symmetrical to the point $MIS = 0.5$ and the performance is the worst at $MIS = 0.5$.

A matched filter can achieve the optimum performance if it can remove the system ISI. When the timing misalignment of CCI is introduced to the system, even though a matched filter can fully remove the ISI caused by the main channel signal, the matched filter can not remove the ISI caused by the co-channel signal since the matched filter is very sensitive to the sampling time. The ISI increases when the timing misalignment of CCI increases. At $MIS = 0.5$, the system ISI reaches its highest point, because in this case the actual shifting of main channel and co-channel is the biggest. From Figure 4.3, we can also notice that the performance degrades significantly when the MIS is in the range of $[0, 0.25]$ and $[0.75, 1.0]$ and that relatively slight changes occur the remaining central region $[0.25, 0.75]$.

Since the SER performance of a matched filter as a function of the MIS is symmetrical to the point $MIS = 0.5$, we only study the system performance assuming that the MIS is in the range of $[0, 0.5]$ in the following simulations. Figure 4.4 shows the simulation result given different MIS as a function of SNR and SIR. In Figure 4.4 [a], the SER performance degradation is small between $MIS = 0.0$ and $MIS = 0.5$, about 3 dB in SNR, under high noise condition. However, for large values of SNR, the performance degradation can be 10 dB in SNR between $MIS = 0.0$ and $MIS = 0.5$. From Figure 4.4 [b], we can get the same conclusion as the Figure 4.4 [a]. The performance degradation which is caused by timing misalignment is more clearly shown in Figure 4.5. From Figure 4.5, we can easily see that the performance degradation is very small when the noise and interference is high (e.g., $SNR < 12.0$ dB for $SIR = 6.0$ dB and $SIR < 2.0$ dB for $SNR = 20.0$ dB), while the performance degradation is relative large when noise and interference is low. We also see that the performance degradations mostly happen in the range of $[0, 0.25]$ which confirms the result obtained from Figure 4.3.

From the above discussions and simulation results, we can see that a matched filter can achieve the optimum SER performance if the main channel and the co-channel are perfectly synchronized. Therefore, it is not necessary to have any kind of equalizer

Figure 4.4 Symbol error rate of a matched filter for different MIS as a function of [a] SNR, (SIR = 6dB) and [b] SIR, (SNR = 20dB)

[a]

[b]

Figure 4.5 Symbol error rate of a matched filter as a function of the MIS at [a] SIR = 6dB and [b] SNR = 20dB

at the receiver end. However, if timing misalignment MIS between the main channel and co-channel exists, the SER performance of a matched filter degrades tremendously, especially when $MIS > 0.1$. A equalizer will be needed to combat this impairment.

4.4 System Performance with Equalizer

4.4.1 System Performance with LMSDFE

Decision feedback equalization is a technique used in digital communication systems to equalize the channel to remove that part of the ISI caused by the previous data decisions. In our simulation, the LMS algorithm is used to adjust the weights of the LMSDFE.

4.4.1.1 Basic Simulation Parameters

For a decision feedback equalizer, the number of forward taps N_f and the number of feedback taps N_b are two very important simulation parameters. The required number of taps in an equalizer to achieve a certain performance is highly dependent on the type of impairment that the equalizer is used for [27, 23]. As we mentioned in chapter 3, the LMS step size, μ , which affects the convergence of the LMSDFE can be considered as another important simulation parameter.

In our system, with a given receiver structure (N_f, N_b) and LMS step size, μ , the system performance is mainly a function of the SNR and SIR at the input to the equalizer. The SNR and SIR, in decibels (dB), have been defined in Equation (3.21) and Equation (3.12). Finally, the normalized timing misalignment of CCI, MIS , can also be a simulation parameter.

Table 4.1 summarizes all the fixed and variable simulation parameters.

4.4.1.2 Convergence Rate

The rate of convergence of LMS for a DFE is investigated in terms of the time evolution of the mean square error (MSE) defined in equation (4.1) and the MSE is performed over an ensemble of 200 trials. In our simulation, the LMS step size, μ ,

Table 4.1 Basic simulation parameters of LMSDFE

(N_f, N_b)	N_f forward and N_b backward taps
μ	LMS Step Size
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
MIS	Timing misalignment of CCI

is fixed to a relative large value during training period, then switched to a small value during data transmit mode to keep tracking the channel.

In this section, we assume that the timing misalignment of CCI, *MIS*, is equal to zero. From previous work, we know that a matched filter can give the optimum performance under this assumption. However, in order to investigate the ability of combating the timing misalignment of CCI, a simple case can be studied first to understand the properties of DFE. In our simulation model, a DFE is used at the receiver end and the additive noise is a colored noise which is generated by passing a AWGN through a bandlimited filter. The DFE structure is chosen as LMSDFE(6, 2) and its structure is depicted in Figure 3.6. Please note that even though this structure may not be the optimum one, it is good enough to help us to see the convergence rate of a LMSDFE using different LMS step size, μ . How to select the optimum structure will be discussed in a later section. Two simulations are done under a higher interference (e.g., *SIR* = 3.0 dB) and a relative lower interference (e.g., *SIR* = 6.0 dB), and the simulation parameters are listed in Table 4.2.

The time evolution of mean square error of the first experiment is plotted in Figure 4.6 [a] for three different step size μ 's, with the *SIR* value at 6.0 dB and SNR at 20.0 dB. The required number of iterations for convergence is defined as the number of interactions required to reach a "flat" slope squared error level. The number of iterations varies for different step size μ 's are shown in Table 4.3. It is seen that the larger the

Table 4.2 LMSDFE convergence rate simulation parameters

	exprimment #1	exprimment #2
# desired paths	1	1
(N_f, N_b)	(6, 2)	(6, 2)
SIR	6 dB	3dB
SNR	20dB	20dB
MIS	0.0	0.0
frame length	2000	2000
simulation trials	200	200

Table 4.3 LMSDFE(6, 2) convergence rate at SIR = 6dB and SNR = 20dB

LMSDFE step size, μ	training length	# iterations for convergence
0.0005/0.0001	1500	1500
0.002/0.0001	600	400
0.02/0.002	200	100

step size μ , the faster the convergence rate. For a LMSDFE with SIR at 6.0 dB, a step size of 0.02/0.002 can produce a reasonably fast rate of convergence.

The rate of convergence of LMSDFE(6, 2) is also tested in a second experiment for a lower SIR value (e.g., 3.0 dB). Comparing Figure 4.6 [a] (SIR = 6 dB) with Figure 4.6 [b] (SIR = 3 dB), it can be seen that the rate of convergence is independent on the values of SIR and SNR. However, with a lower SIR, the LMSDFE is very sensitive to the LMS step size, μ . For example, with step size, $\mu = 0.02/0.002$, the LMSDFE diverges almost immediately after training for 200 symbol periods. It can be concluded that with a relatively large LMS step size, μ , the equalizer's decision-directed adaptation is quite sensitive to error propagation which is more likely to occur with significant interference, e.g., in the case of SIR = 3.0 dB. Therefore, with SIR = 3.0

Figure 4.6 Time evolution of a LMSDFE MSE for [a] SIR = 6 dB SNR= 20dB and [b] SIR = 3 dB SNR = 20dB, MIS = 0.0

dB, the LMSDFE choosing a small step size, $\mu = 0.002$, during training period and $\mu = 0.0001$ during data transmission can provide a reasonably good performance. In this case, comparing with selecting a larger step size, although it converges slower in the first 400 symbol periods, it does maintain this convergence during the data transmission since it is more immune to decision-directed error propagation.

4.4.1.3 Equalizer Length

Although one would expect that an infinite length equalizer is the best equalizer that might be used, it has its own drawbacks. For infinite length DFE, the effect of decision errors, usually called error propagation, may be considerable while this effect is reduced when the number of feedback taps is limited. Moreover, increasing the number of the forward taps beyond some levels, would usually result in a very negligible improvement of the system performance while it makes the equalizer more complex. In this section, we investigate the performance of decision-feedback equalizer with varying structures.

The comparison of equalizers with different forward and backward tap size is shown in Figure 4.7. In Figure 4.7 [a], the experiment is based on a low interference (e.g., *SIR* = 6.0*dB*). The LMS step size, μ , is fixed at 0.02 during training period then switched to 0.002 during data transmit mode and training length is selected as 200 symbols. This plot indicates that the LMSDFE(4, 2) receiver can remove most of the ISI and that an equalizer with longer forward taps (e.g., LMSDFE(6, 2) or LMSDFE(8, 2)) does not improve the system performance at all values of SNR. The slightly poorer performance of the LMSDFE with a large number of forward taps is due to the fact that a longer equalizer, with more tap coefficients, experiences more self-noise in the computation of its tap coefficients. This argument applies for an equalizer with more feed-forward or feedback taps.

The experiment with a high interference (e.g., *SIR* = 3.0 *dB*) is shown in Figure 4.7 [b], where the LMS step size, μ , is 0.002 during training period and 0.0001 during data mode and the training length is 600 symbols. From this plot, we can see that an equalizer with taps size (4, 2) fails to work well at all values of SNR, especially

Figure 4.7 Symbol error rate of a LMSDFE as a function of SNR and DFE configuration at [a] SIR = 6.0 dB [b] SIR = 3.0 dB

at the values of high SNR (e.g., $SNR \geq 16dB$) comparing with LMSDFE(6, 2) or MLMSDFE(8, 2). The reason for this is that when the system distortion due to the high interference is large, a short length of DFE can not remove the ISI introduced by the system. From this figure, we can also see that the performance of the LMSDFE(6, 2) is slight better than that of the LMSDFE(8, 2). Based on these simulation results and considering other limitations specially the convergence properties, a LMSDFE(6, 2) is chosen for the rest of simulation in this thesis.

4.4.1.4 SER Performance without Timing Misalignment of CCI

In the previous two sections, the convergence rate and configuration of LMSDFE have been investigated. We found that the performance of the LMSDFE in terms of mean square error is very sensitive to the choice of the LMS step size under the condition of high interference. Figure 4.8 shows the effect of LMS step size, μ , on the performance of LMSDFE in terms of symbol error rate. From the figure, with lower co-channel interference (e.g., $SIR = 6.0 dB$) the SER performance of LMSDFE using LMS step size $\mu = 0.02/0.002$ and LMSDFE with $\mu = 0.002/0.0001$ is more or less the same. This means that the LMSDFE is insensitive to the LMS step size under condition of low interference which agrees the conclusion from Figure 4.6. However, with higher interference (e.g., $SIR = 3.0 dB$ or $SIR = 2.0 dB$), the SER performance of LMSDFE changes significantly with different values of μ . For instance, when $SIR = 3.0 dB$, the SER performance using step size 0.002/0.001 is much better than using step size 0.02/0.002, especially at the large value of SNR. This plot further confirms the conclusion obtained from the previous simulations.

We also investigate the effect of co-channel interference on the performance of LMSDFE in terms of symbol error rate. The comparison of SER performance of the LMSDFE(6, 2) with a matched filter for different values of SIR is shown in Figure 4.9. From figure 4.9 [a], it is observed that the LMSDFE is quite comparable with a matched filter when $SIR \geq 3.0dB$. It means that a LMSDFE can reach almost optimum SER performance under higher values of SIR. However, for lower values

Figure 4.8 Symbol error rate of a LMSDFE(6, 2) as a function of SNR and LMS step size, μ

Figure 4.9 Symbol error rate of a LMSDFE(6, 2)/a matched filter as a function of SNR

of SIR (e.g., $SIR \leq 2.0dB$), Figure 4.9 [b] shows that the LMSDFE's performance degrades comparing with a matched filter, especially for the large values of $SNR > 20.0 dB$. In general, the LMSDFE can give a very good SER performance under higher values of SIR.

4.4.1.5 SER Performance with Timing Misalignment of CCI

According to the investigation of the SER performance of a matched filter, we made the conclusion that the matched filter can achieve the optimum performance when the main channel and co-channel signals are perfectly synchronized. On the other hand, it does not work well when the timing misalignment of CCI exists since it can not remove the system ISI completely. In this case, an equalizer is needed to combat the system ISI. In this section, a LMSDFE is employed to try to achieve the reasonable SER performance.

A decision feedback equalizer with (6, 2) structure is chosen for the simulation studies. Here, we only study a low interference case, e.g., $SIR = 6.0 dB$. As we find in the previous simulations, the LMS step size, $\mu = 0.02/0.002$, can provide good SER performance and fast convergence rate under condition of low interference. However, when the timing misalignment is introduced into the system, the system distortion is very large, the LMSDFE is very sensitive the large LMS step size even though the interference is low. Therefore, in our simulation, the LMS step size, μ , is selected as 0.002 during training period and 0.0001 during data transmission mode and the training length is 600 symbols. The simulation result is shown in Figure 4.10. From this figure, the SER performance of LMSDFE degrades when the timing misalignment MIS increases. Figure 4.11 shows the comparison of the performance of a LMSDFE(6, 2) and a matched filter. It indicates that the LMSDFE slightly improves the SER performance when the timing misalignment is present in the system.

From these two figures, the LMSDFE does not improve the SER performance very much under timing misalignment of CCI condition. In order to further improve the system performance under these situations, we will investigate the performance of the

Figure 4.10 Symbol error rate of a LMSDFE(6, 2) with different MIS as a function of SNR

Figure 4.11 Symbol error rate of a LMSDFE(6, 2)/a matched filter as a function of timing misalignment, MIS

Table 4.4 Basic simulation parameters of MLPDFE

(N_f, N_b)	N_f forward and N_b backward taps
N_1	Number of Neurons in the first hidden layer
β	BP Learning Parameter
β_i	initial BP Learning Parameter
β_f	final BP Learning Parameter
α	BP Momentum Parameter
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
MIS	Timing misalignment of CCI

MLPDFE in the next section.

4.4.2 System Performance with MLPDFE

The configuration of the multilayer perceptron-based equalizer used in our simulation is a two-layer perceptron and is denoted as $\text{MLPDFE}((N_f, N_b), N_1, I)$. The structure of the MLPDFE is shown in Figure 3.7. A complex-valued backpropagation learning algorithm is employed to adjust the weight coefficients.

4.4.2.1 Basic Simulation Parameters

In our simulation study, the configuration of the MLPDFE, e.g., N_f , N_b , and N_1 is the major simulation parameter. In most of our simulations, the N_f is chosen as $6 \frac{T}{2}$ -spaced taps, while N_b as $2 T$ -spaced taps. These parameters are chosen for the fairness of comparison between the LMSDFE and the MLPDFE. Similar to LMSDFE's step size μ , the learning parameter, β , and momentum parameter, α , are two very important simulation parameters which affect the convergence rate of MLPDFE. The learning and momentum parameters are typically constant in time. However, in our simulation, a variable learning parameter is employed in order to speed up convergence. A "gear-

shifting” algorithm is used to gradually reduce the size of the learning parameter over the training period. Specifically, an initial value of the learning parameter, β_i , is reduced linearly to a final value, β_f , in time steps of N_β symbol periods over a training duration of N_t symbol periods:

$$\beta(nT) = \beta_i - (\beta_i - \beta_f) \left(\frac{nT}{N_t T} \right) \quad n = 0, N_\beta, \dots, N_t \quad (4.14)$$

The momentum parameter is set to a relative large value during training, then set to zero after training.

The system SNR and SIR, as well as timing misalignment of CCI, MIS , can be considered as the basic simulation parameters. All the basic simulation parameters are shown in Table 4.4

4.4.2.2 BP Parameters and Configuration

The BP parameter values and MLP configurations used in our SER simulation study are obtained from preliminary simulations for different values of the BP learning and momentum parameters and MLP configuration. In this preliminary simulation experiment, the performance criterion averaged mean square error, $\overline{mse}$, as defined in equation 4.2, is used to select the approximately optimum BP parameters and MLP configuration. Ideally, these parameter values should optimize the system SER, but in light of the huge amount of simulation time required, the $\overline{mse}$ performance criterion is used instead to quickly obtain some suitable MLPDFE parameters for subsequent SER simulation studies. An MLPDFE((6, 2), N_I , 1) structure is used and results are obtained for the cases of $SIR = 6.0 \text{ dB}$ $SNR = 20.0 \text{ dB}$ and $SIR = 3.0 \text{ dB}$ $SNR = 20.0 \text{ dB}$. The simulation parameters used for these two experiments are listed in Table 4.5.

In order to select the optimum combination of the BP parameters, two simulation experiments are done at $SIR = 6.0 \text{ dB}$ $SNR = 20.0 \text{ dB}$ and $SIR = 3.0 \text{ dB}$ $SNR = 20.0 \text{ dB}$. We first select the BP momentum parameter, α . Figure 4.12 [a] and Figure 4.12 [b] show that a small momentum parameter value of α (e.g., $\alpha \leq 0.5$ for $SIR = 6.0 \text{ dB}$ and $\alpha \leq 0.4$ for $SIR = 3.0 \text{ dB}$) yields better performance than a large value. The reason for this is that during the training period, a large value of α can speed up the convergence rate,

Table 4.5 MLPDFE BP parameters and configuration simulation parameters

	exprimment #1	exprimment #2
# desired paths	1	1
(N_f, N_b)	(6, 2)	(6, 2)
SIR	6 dB	3dB
SNR	20dB	20dB
MIS	0.0	0.0
frame length	2000	2000
simulation trials	10	10

but it can also cause the wrong direction with the possibility of convergence to a local minimum. On the other hand, a small value can smoothly converge to the optimum point. Therefore, in our simulation, the optimum momentum parameter is set as 0.3 or 0.4.

For the case of BP with a gear-shifted learning parameter, β , over the training period (see equation (4.14)), with an optimum momentum value of $\alpha = 0.3$ or 0.4 , there still remain the selection of the optimum value of β_i and β_f . Figure 4.13 shows that the optimum value of β_i is 0.4 or 0.5 for $SIR = 6.0$ dB and 0.4 for $SIR = 3.0$ dB. If the value of β_i is too small, then the MLPDFE weight coefficients converge too slowly or not at all. On the other hand, if the value of β_i is too large, then the adaptation changes in the LMPDFE weight coefficients are too large and the MLPDFE can not converge. The performance of MLPDFE is quite sensitive to the value of β_i with higher interference comparing with lower interference. In our simulation, the optimum BP initial learning parameter, β_i , is selected to be 0.4.

So far, the BP momentum and initial learning parameters have been selected. Based on the selection, Figure 4.14 is the result for selecting the final learning parameter, β_f . The figure shows that the equalizer performance is quite good for most of β_f except for the large values of β_f . If the value of β_f is large, e.g., $\beta_f > 0.05$, then the MLP weight

Figure 4.12 MLPDFE $\overline{mse}$ as a function of momentum parameter, α , at
[a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB

Figure 4.13 MLPDFE $\overline{mse}$ as a function of Initial BP learning parameter, β_{i1} , at [a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB

Figure 4.14 MLPDFE $\overline{mse}$ as a function of final BP learning parameter, β_f at [a] SIR = 6.0 dB SNR = 20.0 dB and [b] SIR = 3.0 dB SNR = 20.0 dB

Figure 4.15 MLPDFE $\overline{mse}$ as a function of number of neurons in the first hidden layer, N_1 , at SIR = 3.0 dB SNR = 20.0 dB

coefficients exhibit significant fluctuations and are thus sensitive to decision errors during adaptations. For $\beta_f < 0.02$, there is no additional performance gain by further reducing β_f . Thus, the β_f is selected as 0.02.

Based on the simulation results, the BP optimum parameters have been selected. The other very important parameter is the number of neurons in the first hidden layer. The MLPDFE performance will vary greatly with different values of N_I . Figure 4.15 shows the effect of choice of N_I on the performance. From the figure, 6 and 15 can have similar MSE performance, but 15 is chosen because Lippmann [12] suggested the number of neurons in the first hidden layer had better be twice of the number of neurons in the input layer. If the value of N_I is too small, the MLPDFE can not learn the received signal completely and can not form the required complex decision region.

4.4.2.3 Convergence Rate

The convergence rate of the MLPDFE is simulated for two cases: one is under a low interference (e.g., $SIR = 6.0$ dB $SNR = 20.0$ dB), the other is under a high interference (e.g., $SIR = 3.0$ dB $SNR = 20.0$ dB). Figure 4.16 shows the convergence performance of MLPDFE((6, 2), 15, 1) (which is the averaged value over 200 trials). Here, the BP learning parameter, β , starts from 0.4 and decreases to 0.005 based on the equation (4.14) during training period. Unlike the LMSDFE whose convergence rate is independent of the values of SIR and SNR and is only affected by the LMS step size, the MLPDFE's convergence rate is not only dependant on the BP parameters and configuration, but also on the values of SIR and SNR. From Figure 4.16, the MLPDFE converges at about 800 symbols under low interference condition while 1500 symbols under high interference condition even though these two cases use the same BP parameters and configuration. From this figure, we can see the MLPDFE seems to be stuck in a local minimum for the initial 300 symbols periods when the interference is high. Figure 4.16 also shows the comparison of the convergence rate of the MLPDFE and the LMSDFE. It is observed that the MSE at the MLPDFE output can be significantly below the channel noise and co-channel interference floor because of its nonlinear processing. Due to the scaling

Figure 4.16 Time evolution of MLPDFE/LMSDFE MSE

and compression of the nonlinear sigmoid function, the MLPDFE output is confined to within the unit variance of the transmitted data symbols and is not linearly related to the channel noise and CCI variance, unlike the case for the LMSDFE with linear processing. On the other hand, because the LMSDFE does not have a compressed nonlinear output, its output MSE is limited by the noise and CCI. Furthermore, the fluctuation in the MSE at the LMSDFE output are commensurately smaller.

4.4.2.4 SER Performance without Timing Misalignment of CCI

In this section, the SER performance of MLPDFE is investigated when there is no time misalignment between main channel and co-channel to see whether MLPDFE is a good approach to combat noise and CCI or not. Here, we employ two different training methods. One is that during both training and data transmission period, the data contain same statistics property, e.g., they have same values of MIS, SIR and SNR. So the received data frame consists of both training data and information data. We call this method as on-line training. The other is that before receiving data, the MLPDFE has been trained with smaller values of SIR and SNR. When the received data frame comes (with any kinds of values of SIR and SNR), the equalizer does not need to be trained any more, but a small BP learning parameter, β , is used to keep tracking the channel. Therefore, the received data frame only contains information data. We call this method off-line training.

Figure 4.17 [a] shows the result of SER of the MLPDFE using on-line training method. The BP learning parameter, β , is selected from 0.4 to 0.02 during the training period and then fixed at 0.02 during the transmission mode. The momentum parameter, α , is chosen as 0.3 during training and 0.0 during data transmission periods. The training length is selected as 1000 symbols for $SIR = 6.0$ dB and 1500 symbols for the rest of values of SIR in the simulation. The reason for choosing longer training length for high interference is that it takes longer to converge. The configuration is MLPDFE((6, 2), 15, 1). From the figure, we can see, just like the performance of a matched filter and a LMSDFE, the increase of CCI degrades the system performance.

Figure 4.17 Symbol error rate of an MLPDFE((6, 2), 15, 1) as a function of SNR using [a] on-line training and [b] off-line training

Figure 4.13 Symbol error rate of an MLPDFE((6, 2), 15, 1) as a function of SNR, comparing the on-line and off-line training

Figure 4.17 [b] shows the result of SER of MLPDFE using off-line training method. The network has been trained using the signal with $SIR = 2.0$ dB and $SNR = 20.0$ dB where the training length is selected as 2000 symbols. The training is done in two steps: during the first 1000 symbol training period, the BP learning parameter, β , starts from the initial value of 0.4 and decreases to 0.002, while the momentum parameter, α , is chosen to be 0.3. In the second 1000 symbol training period, the network is further trained by using $\beta = 0.02$, $\alpha = 0.0$. Comparing these two results using different training methods, we find that the system performance using off-line training is little better than using on-line training, especially at the large values of SNR when the interference is high. Figure 4.17 clearly shows the difference.

In order to illustrate the performance of MLPDFE, a SER performance comparison of a matched filter and a MLPDFE is made assuming that the main channel and co-channel are perfectly synchronized. Figure 4.19 and Figure 4.20 show the comparison of the MLPDFE((6, 2), 15, 1) and a matched filter. Figure 4.19 is the comparison of a MLPDFE using on-line training method and a matched filter, we find that the performance of MLPDFE is very close to the performance of a matched filter. Figure 4.20 is the comparison of MLPDFE employing off-line training method and a matched filter. From this figure, it is observed that the SER performance of the MLPDFE is so close to the performance of a matched filter that we can not tell the difference.

4.4.2.5 SER Performance with Timing Misalignment of CCI

When the main channel and the co-channel are not perfectly synchronized, a matched filter can not work well, especially when the normalized CCI timing misalignment, $MIS > 0.1$. A decision-feedback equalizer has been suggested to be used after the matched filter in order to improve the system performance. From previous simulations and discussions, the LMSDFE only slightly improves the system performance compared to the matched filter, and it still can not work well when the CCI timing misalignment exists. Now, we propose another technique to solve this problem, that is an MLPDFE is employed in our system. In the last section, the properties of MLPDFE have been

Figure 4.19 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training / a matched filter as a function of [a] SNR [a] and [b] SIR

Figure 4.20 Symbol error rate of MLPDFE((6, 2), 15, 1) using off-line training / a matched filter as a function of [a] SNR [a] and [b] SIR

studied in terms of choosing optimum BP parameters and configuration. Two training methods have been discussed and the comparison of the SER performance of these two methods has been made. A comparison of the MLPDFE and a matched filter is also made under the assumption that the main channel and co-channel are synchronized. The simulation results indicate that the performance of the MLPDFE((6, 2), 15, 1) is very comparable to the performance of a matched filter.

In this section, we emphasize the SER performance of an MLPDFE when the main channel and co-channel are not synchronized, that is assuming CCI timing misalignment exists. First of all, a on-line training method is used to investigate the performance of an MLPDFE. Two simulation experiments are done based on a signal with $SIR = 6.0$ dB and a signal with $SIR = 4.0$ dB. Figure 4.21 is the simulation result of the first experiment at $SIR = 6.0$ dB considering different values of CCI timing misalignment. Here, for all different values of CCI timing misalignment, MIS , an MLPDFE((6, 2), 15, 1) equalizer is used, and the training length is selected as 1500 symbols, while the BP learning, $\beta = 0.4 \rightarrow 0.02$, and momentum parameters, $\alpha = 0.3$, during training period. During data transmission period, the β is fixed at 0.02. From this Figure 4.21 [a], it is observed that the SER performance degrades when the CCI timing misalignment increases. There is 3.0 dB of SNR difference between $MIS = 0.0$ and $MIS = 0.5$ at $SER = 10^{-5}$. Figure 4.22 is the result of the second experiment at $SIR = 4.0$ dB. The BP parameters and configuration used here are the same as in the first experiment except that the training length is selected as 3000 symbols. From this figure, we can see the effect of CCI timing misalignment is bigger for higher interference than lower interference. The difference is about 6.0 dB of SNR between $MIS = 0.0$ and $MIS = 0.5$ at $SER = 5 \times 10^{-3}$. Also from Figure 4.21 and Figure 4.22, the degradations mostly happen at large values of SNR.

Figure 4.23 is the simulation result of MLPDFE((6, 2), 15, 1) using off-line training method. Here, we assume that the training data and transmission data must have the same CCI timing misalignment, MIS . This means that if the network is trained by a sequence of training data with a certain value of MIS , then this trained network can only be used to equalize the received signal which has the same CCI timing misalignment.

Figure 4.21 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training as a function of [a] SNR and [b] MIS at SIR = 6.0 dB

[a]

[b]

Figure 4.22 Symbol error rate of MLPDFE((6, 2), 15, 1) using on-line training as a function of [a] SNR and [b] MIS at SIR = 4.0 dB

[a]

[b]

Figure 4.23 Symbol error rate of MLPDFE(6, 2), 15, 1) using off-line training as a function of MIS at [a] SIR = 6.0 dB and [b] SNR = 20.0 dB

2 of/de 2

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

So for the received signal with a certain CCI timing misalignment, we switch it to a network trained by the training data which has a certain value of MIS to handle it. In our simulation, all the networks which are based on certain value of MIS are trained by training signal with $SIR = 2.0\text{ dB}$ and $SNR = 20.0\text{ dB}$. In this simulation, since the interference is quite high, the equalizers need much longer to converge. From this figure, it is observed that the effect of CCI timing misalignment is quite small. Regardless of the CCI timing misalignment, the MLPDFE((6, 2), 15, 1) can give a quite good SER performance.

4.5 Comparison of SER Performance of Three Receivers

4.5.1 SER Performance without Timing Misalignment of CCI

In previous sections of this chapter, the SER performance of a matched filter, a LMSDFE and an MLPDFE have been investigated separately under the assumption that the system has no CCI timing misalignment. Here, the comparison of these three receivers is made under this condition. Figure 4.24 shows the comparison, where Figure 4.24 [a] is the simulation result at $SIR = 6.0\text{ dB}$ and Figure 4.24 [b] is the simulation result at $SIR = 2.0\text{ dB}$. From this figure, the performance of these three receivers are quite close when the interference is low (e.g., $SIR = 6.0\text{ dB}$). This means that the LMSDFE and MLPDFE can almost achieve the optimum performance. On the other, by looking at Figure 4.24 [b], the performance of LMSDFE is considerably worse than the matched filter when the interference is high. However, the MLPDFE can still achieve very good performance when compared with the matched filter.

4.5.2 SER Performance with Timing Misalignment of CCI

In this section, the comparison among these three receivers is made under assumption that CCI timing misalignment exists in the system. Figure 4.25 shows the performance comparison of these three receivers. From this figure, we can see that the increasing of the CCI timing misalignment degrades the performance of these three receivers. The

Figure 4.24 SER performance comparisons among three receivers as a function of SNR at [a] SIR = 6.0 dB and [b] SIR = 2.0 dB

system used a LMSDFE at the receiver end gains slight SER performance improvement comparing using a matched filter. However, the MLPDFE does improve the system performance, especially when the CCI timing misalignment is small (e.g., $MIS < 0.15$). Also the MLPDFE using off-line training method can have much better performance than using on-line training.

[a]

[b]

Figure 4.25 SER performance comparisons among the three receivers as a function of MIS at SIR = 6.0 dB [a] SNR = 16.0 dB and [b] SNR = 18.0 dB

Chapter 5

Computer Simulation results of Adjacent-channel Interference

In this chapter, the SER performance of three receivers, namely a matched filter receiver, a LMSDFE receiver and an MLPDFE receiver is studied in a white gaussian noise channel corrupted by adjacent-channel interference.

5.1 System Model

In this chapter, we use the same system model described in Chapter 3 and consider only two immediate adjacent channels on the left and right hand side of the main channel in the frequency domain as the source for ACI. The effect of the other adjacent channels is considered to be negligible. We assume that the main channel is centered at 0 Hz , and the two immediate adjacent channels are symmetrically located with respect to the main channel and δf is the frequency separation between the main and adjacent channels. To make some quantitative comparisons, the case of only one adjacent channel interference in the system is also considered and the results are given respectively.

5.2 SER Performance with a Matched Filter

We begin to investigate the results of our simulations first by looking at the system performance when no equalizer is used at the receiver, i.e., only a matched filter receiver is used. Figure 5.1 shows the SER curves for different channel spacing. Here, two immediate ACIs which have the same channel spacing respect with main channel are present in the system and their magnitude are equal to that the main channel's signal. These results show that when there is no equalizer in the receiver, especially when the channel spacing is relatively small and the SNR is high, the symbol error is solely determined by SIR.

SIR = 3.64 dB $\delta f T = 0.75$
 SIR = 4.9 dB $\delta f T = 0.80$
 SIR = 6.38 dB $\delta f T = 0.85$
 SIR = 8.2 dB $\delta f T = 0.90$
 SIR = 10.3 dB $\delta f T = 0.95$
 SIR = 13.0 dB $\delta f T = 1.0$

**Figure 5.1 Symbol error rate of a matched filter
for different channel spacing with two ACI**

5.3 SER Performance with Equalizer

In Golestaneh's master thesis studies [24], a decision feedback equalizer was used as a countermeasure to combat excessive adjacent channel interference. Both one interferer and two interferers were considered in his thesis. Square root raised cosine filters were used as transmit and receive filters. Based on the results from the MMSE calculation of DFE in presence of ACI, a DFE with $9 \frac{T}{2}$ -spaced forward taps and 4 T -spaced feedback taps was chosen. Through his simulation study, it was found that the system SER performance was improved dramatically due to the implementation of the LMSDFE(9, 4) and it was more evident for higher SNR. Moreover, by examining the equalizer SER performance, it was found that in all cases, the equalizer could suppress one ACI better than two ACIs for the same SIR and channel spacing. It was also found that for the same SIR, a large channel spacing always led to a better system performance. In the rest of this section, we will look at the system SER performance by using a multilayer perceptron-based decision feedback equalizer and compare it against that of the decision feedback equalizer.

5.3.1 SER Performance with Two ACIs

In this section, two immediate adjacent channels are present in the system and a two-layer perceptron-based equalizer is used at the receiver. Since the ACI timing misalignment had no effect on error rate performance of the system according to Golestaneh's study, only the case of no ACI timing misalignment will be studied. First of all, the SER performance of the MLPDFE will be investigated, then the comparison of the SER performance of MLPDFE and LMSDFE will be made.

In all the simulations, the training method of MLPDFE will be on-line training. The configuration of the MLPDFE is chosen as MLPDFE((9, 4), 6, 1) based on the average mean square error calculation (MSE), and the BP parameters are selected also based on MSE calculation. The simulation parameters used for all the simulations in this section are listed in Table 5.1

In Figure 5.2 [a] and Figure 5.2 [b], the performance of the system is examined for

Table 5.1 MLPDFE BP parameters and configuration simulation parameters

$T = 1$	Symbol duration
(N_f, N_b)	(9, 4)
N_1	6
α	0.2
$(\beta_1 \rightarrow \beta_f)$	(0.3 $\rightarrow$ 0.005)
training length	1900
τ_1	0.0

fixed SIR and different channel spacing, δf . It is assumed that the left and the right ACI are of equal power. In both figures, it is seen that there are some irreducible error floors, especially for small channel spacing. It is also seen that regardless of the amount of the interference, choosing larger channel spacing always leads to a better performance. However, a comparison between these two figures shows that the gain obtained from large channel spacing is less for higher SIR especially for relatively large SNR. These conclusions are same as those obtained by Golestaneh. [24].

Figure 5.3 through 5.6 show the symbol error rates for fixed channel spacing and different SIR. The observations made earlier about the error floor is much more clearly shown in Figure 5.3 and Figure 5.4 for $\delta f = 0.65$ and $\delta f = 0.70$ respectively. On the other hand, when the channel spacing increases to $\delta f = 0.95$, no errors are observed at $SNR > 25$ dB even when the interference is very high (e.g., $SIR = 3.64$ dB). A comparison between Figure 5.3 and Figure 5.6 indicates that for large channel spacing $\delta f = 0.95$, increasing SIR does not improve the symbol error rate as much as it does for small channel spacing $\delta f = 0.65$. This observation was also obtained in [24] according to the simulations.

Figure 5.7 [a] and [b] show the comparison of the SER performance of an MLPDFE and a LMSDFE. Here, the structure of the LMSDFE is chosen as (9, 4) and the training length as 1000 symbols while the LMS step size is selected as 0.02 during training

[a]

[b]

Figure 5.2 MLPDFE((9,4),6,1) Symbol error rate for different channel spacing at fixed [a] SIR = 3.64 dB and [b] SIR = 8.2 dB

Figure 5.3 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.65$

Figure 5.4 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.70$

Figure 5.5 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.75$

Figure 5.6 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of $\delta f = 0.95$

[a]

[b]

Figure 5.7 MLPDFE/LMSDFE Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$

period and 0.002 during data transmission period. It shows that the SER performance of these two types of equalizers is more or less the same. The performance of the MLPDFE is slightly better than that of the LMSDFE under relative higher noise $SNR < 15.0$ dB. On the other hand, the LMSDFE has better performance when the SNR is high.

5.3.2 SER Performance with One ACI

In this section, we investigate the error rate results when only one ACI is present. As we investigate the SER performance of the MLPDFE when two ACIs are present, the configuration of the MLPDFE in this case is also chosen to be an MLPDFE((9, 4), 6, 1). In all the simulations, the training length is selected as 1500 symbols, and the BP parameter is as follows: the BP momentum, α , is selected as 0.2 during training period, and the BP learning parameter is selected at 0.3 as initial value and gradually decreases to 0.005 during training period.

Figure 5.8 shows the results for fixed SIR and different channel spacing. The results for fixed channel spacing and different SIR are shown in Figure 5.9. It can be seen from all these figures that the SER performance pattern is similar in both one and two ACIs. Specifically it is observed from Figure 5.8 that large channel spacing provides better error rate performance. However, as for the effect of SIR, unlike the case of two ACI in which the impact of SIR for small channel spacing is more than large channel spacing, it can be seen from Figure 5.9 that the impact of SIR on small channel spacing is the same as on large channel spacing.

Figure 5.10 shows the comparison of the SER performance of an MLPDFE and a LMSDFE for fixed channel spacing and different SIR. Here, the LMSDFE is selected as (9, 4), the training length is 500 symbols and the LMS step size is selected as 0.02 during training period and 0.002 during data transmission period. It is seen that the MLPDFE performance is slightly better under large noise (e.g., $SNR < 13$ dB) and worse in the rest of the region regardless of the channel spacing. This conclusion is same as the one observed from the case of two ACI in the previous section.

Comparing the results for the case of one ACI with the results for the case of two

Figure 5.8 MLPDFE((9,4),6,1) Symbol error rate for different channel spacing at fixed [a] SIR = 3.64 dB and [b] SIR = 6.38 dB

Figure 5.9 MLPDFE((9,4),6,1) Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$

[a]

[b]

Figure 5.10 MLPDFE/LMSDFE Symbol error rate for different SIR at fixed channel spacing of [a] $\delta f = 0.65$ and [b] $\delta f = 0.95$

ACIs, it is observed that for the same SIR the system performance with one ACI is much better than that with two ACIs regardless of the channel spacing and the types of equalizers.

Chapter 6

Conclusion and Recommendations

6.1 Conclusion

In this thesis, an equalizer based on a two-layer perceptron structure with complex-valued backpropagation algorithm was used as a countermeasure to combat the inter-symbol interference, additive noise, co-channel interference and adjacent-channel interference.

First of all, the study of the performance of the MLPDFE was made based on the presence of only CCI. The effect of the MLPDFE configuration and BP parameters was investigated in terms of the averaged mean square error. An MLPDFE((6, 2), 9, 1) was chosen as the standard neural network equalizer throughout this section. We studied the convergence rate of the MLPDFE and found that it was much slower than that of the LMSDFE. A system with and without CCI timing misalignment between the main channel and the co-channel were under study and the comparison of the SER performances of the MLPDFE((6, 2), 9, 1), the LMSDFE(6, 2) and a matched filter was made under these two conditions. When the co-channel and main channel are perfectly synchronized, it was found that the matched filter could provide the optimum SER performance, and the MLPDFE((6, 2), 9, 1) could achieve similar performance as the matched filter at all different values of SIR while the LMSDFE(6, 2) could match the performance of the matched filter only at large value of SIR (e.g., $SIR > 3.0\text{ dB}$). On the other hand, when co-channel and main channel are not synchronized, it was found that the SER performance of the matched filter degraded greatly as CCI timing misalignment increased. A LMSDFE(6, 2) did not improve the system performance very much when CCI timing misalignment existed in the system. An MLPDFE((6, 2), 9, 1) was used based on two types of training methods, namely on-line training and off-line training, to combat the CCI. It was observed that the system performance had been improved greatly, especially by using off-line training method.

Following the investigation of the performance of the MLPDFE with CCI, the study was made based on the presence of only ACI. Both one adjacent-channel and two-adjacent channels were considered and the assumption of perfect ACI timing alignment was made. It was found that for the fixed SIR, a large channel spacing always led to a better SER by using an MLPDFE((9, 4), 6, 1) regardless of one ACI or two ACI in the system. On the other hand, the impact of SIR on SER was found to be more severe for smaller channel spacing in two ACI environment while the SIR had almost the same effect on SER for different channel spacing in the case of one ACI environment. It was also observed that in all the case, the MLPDFE could suppress one ACI better than two ACI for the same SIR. The comparison between the SER performance of the MLPDFE((9, 4), 6, 1) and the LMSDFE(9, 4) was made in both cases of one ACI and two ACI. It was found that the MLPDFE((9, 4), 6, 1) had very similar SER performance as that of the LMSDFE(9, 4). Moreover, the MLPDFE((9, 4), 6, 1) could achieve slightly better performance at lower value of SNR (e.g., $SNR < 15$ dB for two ACI and $SNR < 13$ dB for one ACI) while in the case of low noise, the LMSDFE(9, 4) performed a little bit better.

6.2 Recommendations for Future Research

The aim of this thesis was to investigate some of the capabilities of a multilayer perceptron-based decision feedback equalizer in the presence of ISI, additive noise, CCI and ACI. The results presented in this thesis were based on certain assumptions and simulations. Perhaps there are other MLPDFE configurations and BP parameters which can perform even better than that studied in this thesis, especially for the case of ACI environment. In addition, the MLPDFE convergence is very slow and requires improvement. The radial-basis function approach can be employed as a means to improve the adaptation speed of the MLPDFE.

Fading, being a major impairment in all radio system, was not addressed in this thesis. The ability of the multilayer perceptron-based equalizer to track a time-varying channel should be considered for further research.

Appendix A Complex-valued Backpropagation Algorithm

A.1 The Generalized Delta Rule

In this part, we present the formal mathematical description of complex-valued backpropagation algorithm (BP). The learning law used in the BP algorithm is the generalized delta rule (GDR) which is used to minimize the sum of the squares of the errors for all output neurons:

$$\epsilon^2 = \frac{1}{2} \sum_{k=0}^{N_M-1} |E_k^{(M)}|^2 \quad (\text{A.1})$$

where

N_M is the number of neurons in the output layer

$E_k^{(M)}$ is the complex form of the error for the k th neuron in the output layer and

The factor of $\frac{1}{2}$ in equation (A.1) is there for convenience in calculating derivatives later. Since an arbitrary constant will appear in the final result, the presence of this factor does not invalidate the derivation.

A.2 Update of Output Layer Weights

In the derivation of the delta rule, the error for the k th neuron in the output layer is :

$$E_k^{(M)} = D_k - X_k^{(M)} \quad (\text{A.2})$$

where

D_k is the desired output.

$X_k^{(M)}$ is the actual output which can be expressed as:

$$\begin{aligned} X_k^{(M)} &= F(NE T_k^{(M)}) \\ &= F\left(\sum_{j=0}^{N_{M-1}-1} W_{jk}^{(M)} X_j^{(M-1)}\right) \end{aligned} \quad (A.3)$$

where

$W_{jk}^{(M)}$ is the weights in the output layer

N_{M-1} is the number of the neuron in the hidden layer which is next to the output layer

$X_j^{(M-1)}$ is the input to the output layer and

$$F(NE T_k^{(M)}) \quad (A.4)$$

is a complex-valued sigmoid function. In [16], Benvenuto proposed:

$$F(X) = f(\text{Re}[X]) + j f(\text{Im}[X]) \quad (A.5)$$

where X is a complex variable and $f(x)$ is the conventional real-valued sigmoid function shown in equation (2.2). Now $F(X)$ is always bounded and actually its behavior is a 2-D extension of $f(x)$. Moreover, even if $F(X)$ is not analytic, a simple extension of the standard BP algorithm can be found.

Since the BP algorithm is a steepest descent search technique, the weights in the M th layer are adjusted in the opposite direction to the gradient of the squared error at the output:

$$\Delta W_{jk}^{(M)} \propto \nabla_{W_{jk}^{(M)}} e^2 \quad (A.6)$$

where the change in weight coefficient for the n th symbol period is:

$$\Delta W_{jk}^{(M)}(n) = W_{jk}^{(M)}(n+1) - W_{jk}^{(M)}(n) \quad (A.7)$$

and the complex-valued gradient operator is :

$$\nabla_{W_{jk}^{(M)}} = \frac{\partial}{\partial \text{Re}[W_{jk}^{(M)}]} + j \frac{\partial}{\partial \text{Im}[W_{jk}^{(M)}]} \quad (A.8)$$

In order to find a formula for the BP algorithm, it is required to compute the gradient $\nabla_{W_{jk}^{(M)}} e^2$ or, equivalently, the partial derivatives $\frac{\partial e^2}{\partial \text{Re}[W_{jk}^{(M)}]}$ and $\frac{\partial e^2}{\partial \text{Im}[W_{jk}^{(M)}]}$. These partial derivatives are taken with respect to the real and imaginary parts of each weight coefficient.

To determine the direction in which to change the weights, we first calculate the negative of the gradient of e^2 , with respect to $X_k^{(M)}$:

$$\begin{aligned} E_k^{(M)} &= \frac{1}{2} \nabla_{X_k^{(M)}} (e^2) \\ &= -\frac{1}{2} \frac{\partial e^2}{\partial \text{Re}[X_k^{(M)}]} - \frac{j}{2} \frac{\partial e^2}{\partial \text{Im}[X_k^{(M)}]} \end{aligned} \quad (\text{A.9})$$

Similar, we define

$$\begin{aligned} \Delta_k^{(M)} &= -\frac{1}{2} \nabla_{NET_k^{(M)}} (e^2) \\ &= -\frac{1}{2} \frac{\partial e^2}{\partial \text{Re}[NET_k^{(M)}]} - \frac{j}{2} \frac{\partial e^2}{\partial \text{Im}[NET_k^{(M)}]} \end{aligned} \quad (\text{A.10})$$

Based on (A.5), (A.3), and (A.9), it results

$$\begin{aligned} \Delta_k^{(M)} &= -\frac{1}{2} \frac{\partial e^2}{\partial x_{kR}^{(M)}} \frac{\partial x_{kR}^{(M)}}{\partial \text{net}_{kR}^{(M)}} - \frac{j}{2} \frac{\partial e^2}{\partial x_{kl}^{(M)}} \frac{\partial x_{kl}^{(M)}}{\partial \text{net}_{kl}^{(M)}} \\ &= e_{kR}^{(M)} f'(\text{net}_{kR}^{(M)}) + j e_{kl}^{(M)} f'(\text{net}_{kl}^{(M)}) \end{aligned} \quad (\text{A.11})$$

Finally, the gradient of the square error e^2 with respect to the weight $W_{jk}^{(M)}$ is

$$\begin{aligned} \nabla_{W_{jk}^{(M)}} (e^2) &= \frac{\partial e^2}{\partial \text{Re}[W_{jk}^{(M)}]} + j \frac{\partial e^2}{\partial \text{Im}[W_{jk}^{(M)}]} \\ &= \left(\frac{\partial e^2}{\partial \text{Re}[NET_k^{(M)}]} \frac{\partial \text{Re}[NET_k^{(M)}]}{\partial \text{Re}[W_{jk}^{(M)}]} + \frac{\partial e^2}{\partial \text{Im}[NET_k^{(M)}]} \frac{\partial \text{Im}[NET_k^{(M)}]}{\partial \text{Re}[W_{jk}^{(M)}]} \right) \\ &\quad + j \left(\frac{\partial e^2}{\partial \text{Re}[NET_k^{(M)}]} \frac{\partial \text{Re}[NET_k^{(M)}]}{\partial \text{Im}[W_{jk}^{(M)}]} + \frac{\partial e^2}{\partial \text{Im}[NET_k^{(M)}]} \frac{\partial \text{Im}[NET_k^{(M)}]}{\partial \text{Im}[W_{jk}^{(M)}]} \right) \end{aligned} \quad (\text{A.12})$$

Based on (A.3) and (A.10), one obtains

$$\begin{aligned}\nabla_{W_{jk}^{(M)}}(\epsilon^2) &= -2 \left(\text{Re} \left[\delta_k^{(M)} \right] \text{Re} \left[X_k^{(M-1)} \right] + \text{Im} \left[\delta_k^{(M)} \right] \text{Im} \left[X_k^{(M-1)} \right] \right) \\ &\quad - 2j \left(-\text{Re} \left[\delta_k^{(M)} \right] \text{Im} \left[X_k^{(M-1)} \right] + \text{Im} \left[\delta_k^{(M)} \right] \text{Re} \left[X_k^{(M-1)} \right] \right) \quad (\text{A.13}) \\ &= -2 \Delta_k^{(M)} X_k^{*(M-1)}\end{aligned}$$

where

* is the conjugate of the complex.

Thus, the form of the BP weight update equation for the weights in the output layer is:

$$\Delta W_{jk}^{(M)} = \beta \Delta_k^{(M)} X_k^{*(M-1)} \quad (\text{A.14})$$

A.3 Update of the Hidden Layer Weights

In order to adapt the weight coefficients in the hidden layer, the error at the output layer must be backpropagated to these hidden layers. The BP algorithm updates the weights of each layer in succession starting from the output layer and moving back towards the input layer. In order to determine the direction in which to change the weights between neurons in the hidden layer and the input layer, it is required to compute gradient $\nabla_{W_{ij}^{(M-1)}}(\epsilon^2)$ with respect to the real and imaginary parts of the weights in the $(M-1)th$ layer.

For a hidden layer ($0 < l < M$), expression (A.1) can be expanded using the chain rule as follows:

$$\begin{aligned}E_j^{(l)} &= -\frac{1}{2} \sum_{k=0}^{N_l-1} \left(\frac{\partial \epsilon^2}{\partial \text{Re} [NET_k^{(l)}]} \frac{\partial \text{Re} [NET_k^{(l)}]}{\partial \text{Re} [X_j^{(l-1)}]} + \frac{\partial \epsilon^2}{\partial \text{Im} [NET_k^{(l)}]} \frac{\partial \text{Im} [NET_k^{(l)}]}{\partial \text{Re} [X_j^{(l-1)}]} \right) \\ &\quad + j \left(\frac{\partial \epsilon^2}{\partial \text{Re} [NET_k^{(l)}]} \frac{\partial \text{Re} [NET_k^{(l)}]}{\partial \text{Im} [X_j^{(l-1)}]} + \frac{\partial \epsilon^2}{\partial \text{Im} [NET_k^{(l)}]} \frac{\partial \text{Im} [NET_k^{(l)}]}{\partial \text{Im} [X_j^{(l-1)}]} \right) \quad (\text{A.15})\end{aligned}$$

Using definition (A.10) and relationship (A.3) gives

$$\begin{aligned}
 E_j^{(l)} = & \sum_{k=0}^{N_{l+1}-1} \left(\text{Re} \left[\delta_k^{(l+1)} \right] \text{Re} \left[W_{kj}^{(l+1)} \right] + \text{Im} \left[\delta_k^{(l+1)} \right] \text{Im} \left[W_{kj}^{(l+1)} \right] \right) \\
 & + j \sum_{k=0}^{N_{l+1}-1} \left(-\text{Re} \left[\delta_k^{(l+1)} \right] \text{Im} \left[W_{kj}^{(l+1)} \right] + \text{Im} \left[\delta_k^{(l+1)} \right] \text{Re} \left[W_{kj}^{(l+1)} \right] \right)
 \end{aligned} \tag{A.16}$$

As a summary of the BP algorithm, the weight adaptation form is follows:

For $l = M, \dots, 1$ and $j = 0, 1, \dots, N_l - 1$

$$W_{ij}^{(l)}((n+1)T) = W_{ij}^{(l)}(nT) + \beta \Delta_j^{(l)} X_i^{*(l-1)} \tag{A.17}$$

where

$$E_k^{(l)} = \begin{cases} E_k^{(M)}, & \text{for } l = M \\ \sum_{j=1}^{N_{l+1}} W_{jk}^{*(l)} \Delta_j^{(l+1)}, & \text{for } l \neq M \end{cases} \tag{A.18}$$

$$\Delta_k^{(l)} = \text{Re} \left[E_j^{(l)} \right] f' \left(\text{Re} \left[NET_j^{(l-1)} \right] \right) + j \text{Im} \left[E_j^{(l)} \right] f' \left(\text{Im} \left[NET_j^{(l-1)} \right] \right) \tag{A.19}$$

Appendix B Complex-Valued LMS Algorithm

For the case of a DFE (and all linear filters in general), the squared error at its output is a quadratic function of its weights. Thus, there exists a global minimum on the squared error surface. The complex-valued LMS algorithm is based on the stochastic gradient steepest descent search method [adaptive]. That is, the weights are adjusted in the opposite direction to the gradient of the squared error at the output of the DFE so as to achieve the minimum squared error. Thus, the complex-valued LMS weights update equation is:

$$\underline{W}((n+1)T) = \underline{W}(nT) - \frac{\mu(nT)}{2} \nabla_{\underline{W}(nT)} E(nT) \quad (\text{B.1})$$

where the squared error at the DFE output is:

$$E(nT) = \varepsilon(nT)\varepsilon^*(nT) = \left| \hat{a}((n-D)T) - \underline{x}'(nT)\underline{W}(nT) \right|^2 \quad (\text{B.2})$$

and the DFE coefficient vector, $\underline{W}(nT)$, is comprised of the feedforward coefficient vector, $\underline{W}_f(nT)$, and the feedback coefficient vector, $\underline{W}_b(nT)$,

$$\underline{W}(nT) = \begin{bmatrix} \underline{W}_f(nT) \\ \underline{W}_b(nT) \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} W_0(nT) \\ W_1(nT) \\ \vdots \\ W_{N_f-1}(nT) \end{bmatrix} \\ \begin{bmatrix} W_{N_f}(nT) \\ W_{N_f+1}(nT) \\ \vdots \\ W_{N_f+N_b-1}(nT) \end{bmatrix} \end{bmatrix} \quad (\text{B.3})$$

while the input data vector, $\underline{x}(nT)$, consists of the received data vector, $\underline{r}(nT)$, in the feedforward tapped delay line and the decision data (from the decision device output)

vector, $\hat{\underline{a}}(nT)$, in the feedback tapped delay line (recall equation 3.24):

$$\underline{x}(nT) = \begin{bmatrix} \underline{r}(nT) \\ \hat{\underline{a}}(nT) \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} r(nT) \\ r(nT - \frac{T}{2}) \\ \vdots \\ r(nT - (N_f - 1)\frac{T}{2}) \end{bmatrix} \\ \begin{bmatrix} \hat{a}((n - D - 1)T) \\ \hat{a}((n - D - 2)T) \\ \vdots \\ \hat{a}((n - D - N_b)T) \end{bmatrix} \end{bmatrix} \quad (\text{B.4})$$

and $\mu(nT)$ is the LMS step size, D is the delay in the system, $*$ is the complex conjugate operator and $'$ is the vector transpose operator. For a given channel characteristic (specifically), the eigenvalue spread of its autocorrelation matrix, there exists a well-defined range of value for the step size which ensures stable convergence of the LMS algorithm [7]. As a rule of thumb for stable convergence, the LMS step size is approximately the inverse of the product of the signal variance at the input of the equalizer and the number of equalizer taps.

In order to derive the formula for the complex-valued LMS algorithm for the DFE, it is required to compute the gradient $\nabla_{\underline{W}(nT)} E(nT)$. The squared error, $E(nT)$, is real-valued and is not analytic in the complex plane. Consequently, the gradient operator must be performed separately for the real and imaginary parts of the DFE weights:

$$\nabla_{\underline{W}(nT)} E(nT) = \frac{\partial E(nT)}{\partial \text{Re}[\underline{W}(nT)]} + j \frac{\partial E(nT)}{\partial \text{Im}[\underline{W}(nT)]} \quad (\text{B.5})$$

Specifically, the partial derivable with respect to the real part of the weight coefficient vector and be written as:

$$\begin{aligned} \frac{\partial E(nT)}{\partial \text{Re}[\underline{W}(nT)]} &= \frac{\partial \varepsilon(nT)}{\partial \text{Re}[\underline{W}(nT)]} \varepsilon^*(nT) + \varepsilon(nT) \frac{\partial \varepsilon^*(nT)}{\partial \text{Re}[\underline{W}(nT)]} \\ &= -\underline{x}(nT) \varepsilon^*(nT) + \varepsilon(nT) (-\underline{x}^*(nT)) \\ &= -2\text{Re}[\varepsilon(nT) \underline{x}^*(nT)] \end{aligned} \quad (\text{B.6})$$

while the partial derivative with respect to the imaginary part of the weight coefficient vector is:

$$\begin{aligned}
 \frac{\partial E(nT)}{\partial \text{Im}[\underline{W}(nT)]} &= \frac{\partial \varepsilon(nT)}{\partial \text{Im}[\underline{W}(nT)]} \varepsilon^*(n) + \varepsilon(nT) \frac{\partial \varepsilon^*(nT)}{\partial \text{Im}[\underline{W}(nT)]} \\
 &= -j \underline{x}(nT) \varepsilon^*(nT) + \varepsilon(nT) (-j \underline{x}^*(nT)) \\
 &= -2 \text{Im}[\varepsilon(nT) \underline{x}^*(nT)]
 \end{aligned} \tag{B.7}$$

These partial derivatives are taken with respect to each real and imaginary component of the weights vector. Combining the partial derivatives with respect to both the real and imaginary parts of the DFE weight vector in equation B.6 and B.7 yields the complex-valued LMS weight coefficient update equation (recall equation 3.25):

$$\underline{W}((n+1)T) = \underline{W}(nT) + \mu(nT) \varepsilon(n) \underline{x}^*(nT) \tag{B.8}$$

References

- [1] D.C.Cox, "Universal digital portable radio communications," in *Proc. IEEE*, vol. 75, no. 4, pp. 436-477, Apr. 1987.
- [2] P. Cochrane and M. Brain, "Future optical fiber transmission technology and networks," *IEEE Communications Magazine*, vol. 26, no. 26, pp. 45-60, Nov. 1988.
- [3] G. Kaplan, "A special guide to data communications," *IEEE Spectrum*, vol. 28, no. 8, pp.21, Aug. 1991.
- [4] C. L. Jackson, "The allocation of the radio spectrum," *Scientific American*, vol. 242, no. 2, pp. 34-39, 1980.
- [5] D. C. Cox, "Cochannel interference considerations in frequency reuse small-coverage-area radio systems," *IEEE Tran. on Commun.*, vol. COM-30, no. 1, pp. 135-142, Jan. 1982.
- [6] S. Wang and S. S. Rappaport, "Signal-to-interference calculations for balanced channel assignment patterns in cellular communications," *IEEE Trans. on Commun.*, vol. COM-37, no. 10, pp. 1077-1087, Oct. 1989.
- [7] S. Qureshi, "Adaptive equalization," in *Proc. of the IEEE*, vol. 73, no. 9, pp. 1349-1387, Sep. 1985.
- [8] J. Horikoshi, "Error performance improvement of QDPSK in the presence of cochannel or multipath interference using diversity," *Electronics and communications in Japan*, vol. 67-B, No. 4, 1984.
- [9] W. A. Gardner and S. Venkataraman, "Performance of optimum and adaptive frequency-shift filters for co-channel interference and fading," in *twenty-four annual asilomar conference on signals, system and computers*, Asilomar Hotel, Pacific Grove, California, Nov. 5-7, 1990.
- [10] J. G. C. Klimasauska and G. Pelton, *Volume 1, Neural Computing*. NeuralWare Inc., 1989.
- [11] G. H. D. Rumelhart and R. Williams, "Learning internal representations by error propagation," in *Parallel Distribution Processing*, Cambridge, Ma, MIT press, vol. 1,, 1986.
- [12] R. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*., Apr. 1987.
- [13] M. Minsky and S. Papert, *Perceptron: An introduction to computational geomery*. MIT Press, 1988.
- [14] R. Duda and P. Hart, *Pattern classification and scence analysis*. John Wiley and Sons, New York, 1973.
- [15] J. M. B. Window and M. Ball, "The Complex LMS Algorithm," in *in Proc. of the IEEE*, vol. 63, no. 4, pp. 719-720, Apr. 1975.
- [16] N. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *IEEE Tran. on Signal Processing*, vol. 40, no. 4, Apr. 1992.

- [17] C. Klimasauskas, "Neural nets and noise filtering," *Dr. Dobby's Journal*, pp. 32–47, Jan. 1989.
- [18] G. Gibson and S. Siu, "Multilayer perceptron structures applied to adaptive equalizers for data communications," in *IEEE Conf. on Neural Networks*, pp. 1183–1186, 1989.
- [19] G. G. S. siu and C. Cowan, "Decision feedback equalization using neural network structures and performance comparison with standard architecture," in *IEE Proc.* vol. 137, no. 4, pp. 221–225, 1990.
- [20] N. W. K. Lo, "Multilayer Perceptron Equalization in the Presence of Co-channel Interference and Noise." Project report.
- [21] M. Tahernezehadi and J. Cruz, "Frequency Selective Fading and Adjacent Channel Interference Performance of MSK and SFSK," *IEEE Trans. on Communications*, vol. 38, no. 1, pp. 102–108, Jan. 1990.
- [22] P. Nigier and P. Vandamme, "Performance of Equalization Techniques in a Radio Interference Environment," *IEEE Trans. on Communications*, vol. 39, no. 3, pp. 452–457, March 1991.
- [23] K. H. M. Uesugi and K. Tsubaki, "Adaptive equalization in TDMA digital mobile radio," in *Conf. Rec. IEEE Globecom 89*, vol. 1, (Dallas, TX), pp. 95–101, Nov. 27–30, 1989.
- [24] S. Golestaneh, "Adjacent Channel Interference Cancellation by Adaptive Equalizers," Master's thesis, Department of System and Computer Engineering, Carleton University, 1992.
- [25] J. Proakis, *Digital Communications*. NJ: McGraw-Hill Inc., 1989.
- [26] K. Feher, *Digital Communications: Satellite/Earth Station Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [27] M. Despinic and D. P. Taylor, "A Simulation study of ACI and TCM in M-QAM Digital Microwave Radio," in *Faculty of Engineering, McMaster University, Hamilton, Ontario, Canada*, CRL Report No. 196, Oct., 1989.

END

2 1-0 7-9 4

FIN