BIDIRECTIONAL DECISION FEEDBACK EQUALIZATION AND MIMO CHANNEL TRAINING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jaiganesh Balakrishnan

August 2002

© Jaiganesh Balakrishnan 2002

ALL RIGHTS RESERVED

BIDIRECTIONAL DECISION FEEDBACK EQUALIZATION AND MIMO CHANNEL TRAINING

Jaiganesh Balakrishnan, Ph.D.

Cornell University 2002

A major obstacle in reliable digital communication is inter-symbol interference (ISI), which is encountered in transmission over frequency-selective channels. A decision feedback equalizer (DFE) offers an effective and a low-complexity solution to combat ISI. However, the DFE is suboptimal and has a performance gap from the matched filter bound. In addition, the DFE suffers from error propagation caused by the feedback of incorrect decisions.

The increase in popularity of packet based transmission systems like GSM or EDGE offers the possibility of block processing of the received signal. With block processing comes the freedom to process the signal in either a causal or a non-causal fashion. A novel bidirectional decision feedback equalizer (BiDFE) architecture that employs time-reversal of the received block of data is proposed in this dissertation. The BiDFE consists of two parallel DFE structures, one to equalize the received signal and the other to equalize the time-reversed version of the received signal. The BiDFE architecture is shown to provide a significant performance improvement over a conventional DFE with little additional complexity.

To gain insight into the performance limitations of the BiDFE, the asymptotic (as the noise variance approaches zero) mean-squared error (MSE) performance of an infinite length design is evaluated. In an attempt to further improve performance, the filter

coefficients of the BiDFE are optimized to minimize the overall MSE. However, when the ideal feedback assumption is relaxed, the symbol-error-rate (SER) performance does not show an improvement. To overcome this problem, two approaches that offer an additional improvement in SER performance, albeit marginal, are proposed.

The BiDFE architecture is extended to the multiple-input multiple-output (MIMO) channel equalization. The design of the BiDFE assumes knowledge of the channel impulse response, which is typically estimated at the receiver. In training based MIMO channel estimation, the choice of training sequence affects performance. The criteria of optimality of MIMO training sequences is derived and design trade-offs in the choice of training length are discussed.

Biographical Sketch

Jaiganesh Balakrishnan was born in the city of Madras (now renamed Chennai), India in the year 1976. The first 17 years of his life were spent in Chromepet, a suburb of Madras. He did his schooling at N.S.N Matriculation Higher Secondary School, Chromepet. In 1993, he joined the Bachelor of Technology program in Electrical Engineering at the Indian Institute of Technology, Madras.

He entered graduate school at Cornell University, Ithaca, NY, in Fall 1997 and received his M.S degree in electrical engineering in August 1999. In the Summer of 1998 and 2000, he interned with the Wireless Technologies Research Department, Bell Labs, Lucent Technologies at Crawford Hill, NJ. After graduation, he plans to join the Mobile Wireless Research Lab, Texas Instruments at Dallas, TX. His research interests include equalization, detection and estimation, adaptive signal processing and wireless communications.

He is an avid birder and spends most of his spare time watching birds. In the Summer of 2001, he spent a few weeks in an island, off the coast of Maine, working as a research intern for the Seabird Restoration Program with the National Audubon Society. He is the winner of the 2001 McIlroy birding competition, held in Ithaca, NY.

To my sister, Santhi

and my parents,

Vasantha and Balakrishnan

Acknowledgements

I would like to express my gratitude to Rick Johnson for his interest and involvement with my research. His constant input has gone towards improving my technical writing style. I have greatly benefited from the emphasis that Rick places on collaboration with other academic and industrial researchers. He encourages his students to pursue their interests, both academic and non-academic, and this has been a primary factor in making my stay at Cornell enjoyable.

I am grateful to Dr. Lang Tong and Dr. Toby Berger for serving in my committee and for their feedback on my dissertation. I thank Lucent Technologies for giving me an opportunity to do an internship with their wireless technology research group and for their generous gift to my research group, C. U. BERG. I would like to thank Harish Viswanathan and Markus Rupp, both from Lucent Technologies, for their insight and contribution on the problem of training sequence design for MIMO channel estimation. Thanks also for assistance from NSF Grant ECS-9811297, Applied Signal Technology, and Nxtwave Communications. I would like to acknowledge Wonzoo Chung (currently at Dotcast), Rick Martin and Andy Klein for many interesting discussions and the relaxed atmosphere in the office.

Thanks to all of my friends for making me feel at home in Ithaca. Vinayak has been a good friend for the last five years and I thank him for the many illuminating discussions

on topics as varied as politics, sports, society and culture. Ashok has been an inspiration to me and conversations with him are always enlightening. Ganesh, with his dedication to research and Tamil cinema, has been an invaluable friend. Many thanks to Sowmya for being a caring friend and for providing a different perspective on life. Venkat, with his good sense of humor and quick wit, is a delight to be with. Thanks to Anurag for the delicious home-cooked North Indian food.

A significant portion of my spare time during the past three years has been spent on birding. Antony, an ardent lover of parrots, has been instrumental in arousing my enthusiasm for birds. I thank all my birding friends from Ithaca for sharing their knowledge on birds and birding localities. The numerous species of birds that I have seen and enjoyed in the Cayuga lake basin are listed in Appendix C.

I thank Dr. S. Sundaram for his encouragement and advice all through my life and Dilip for being my friend, philosopher and guide. I utilize this opportunity to express my deep gratitude to my parents and sister, to whom this dissertation is dedicated, for their love and affection. They have sacrificed a lot for my sake.

Table of Contents

1 Introduction **1**
- 1.1 Background . . . 2
- 1.2 Motivation . . . 7
- 1.3 Organization . . . 11

2 Bidirectional DFE **14**
- 2.1 System Model . . . 15
  - 2.1.1 Matched Filter Bound . . . 19
- 2.2 DFE Review . . . 20
  - 2.2.1 Finite length MMSE-DFE . . . 23
  - 2.2.2 Infinite length MMSE-DFE . . . 25
  - 2.2.3 Gap from the MFB . . . 26
  - 2.2.4 Error Propagation . . . 27
  - 2.2.5 Numerical Example . . . 28
- 2.3 DFE Enhancements . . . 29
  - 2.3.1 Non-causal DFE . . . 29
  - 2.3.2 Decoder Feedback . . . 32
  - 2.3.3 Decision Device Optimization . . . 33
- 2.4 Time-reversal . . . 33
  - 2.4.1 Selective Time-Reversal DFE . . . 35
- 2.5 Bidirectional DFE . . . 36
  - 2.5.1 Rationale . . . 38
  - 2.5.2 BAD: Bidirectional Arbitrated DFE . . . 38
  - 2.5.3 Preliminaries . . . 39
  - 2.5.4 Diversity Combining . . . 41
  - 2.5.5 Time-Reversal Diversity . . . 43
  - 2.5.6 Simulation Results . . . 44
  - 2.5.7 Implementation Issues . . . 50
- 2.6 Summary . . . 56

3 Infinite Length BiDFE **58**
- 3.1 System Model . . . 59
- 3.2 Performance of an Infinite Length LC-BiDFE . . . 60
  - 3.2.1 Numerical Results . . . 65

3.3 LC-BiDFE Tap Optimization . . . 67
3.3.1 Relation to MSE optimized NCDFE . . . 71
3.3.2 Uniqueness of MMSE-BiDFE . . . 71
3.4 Summary . . . 73

4 Finite Length BiDFE **74**
4.1 MMSE-BiDFE for a Symmetric Channel . . . 75
4.1.1 Numerical Results . . . 78
4.2 LC-BiDFE tap optimization with modified cost . . . 80
4.3 LC-BiDFE with Iteration . . . 84
4.4 MMSE-BiDFE for an Asymmetric Channel . . . 84
4.5 Summary . . . 87

5 BiDFE for MIMO Channel Equalization **88**
5.1 System Model . . . 89
5.1.1 Multichannel Matched Filter Bound . . . 91
5.2 MIMO Equalization . . . 91
5.3 BiDFE Extension to MIMO DFE . . . 94
5.4 Numerical Results . . . 95
5.5 Summary . . . 98

6 Training Sequence Design for MIMO Channel Estimation **100**
6.1 System Model . . . 101
6.2 MIMO Channel Estimation . . . 102
6.3 Training Sequence Design . . . 104
6.3.1 Training Sequence Length Design . . . 106
6.4 Search for good Training Sequences . . . 108
6.4.1 Full search . . . 109
6.4.2 Random Search . . . 110
6.4.3 Cyclic Shift Search . . . 110
6.5 Training Sequence for Delay-diversity Scheme . . . 111
6.6 Summary . . . 115

7 Conclusions **117**
7.1 Summary of Results . . . 117
7.2 Future Directions . . . 119

A Additional Simulation Results **121**
A.1 Broadband Wireless Channels . . . 121
A.2 BiDFE Performance for a Channel with Deep Nulls . . . 123
A.3 Simulation Example for LC-BiDFE with Iteration . . . 125
A.4 Fading Channel Simulation for MIMO LC-BiDFE . . . 126

B Training Sequences for EDGE **129**

C Birding the Cayuga Lake Basin **134**

Bibliography **139**

List of Tables

2.1 Known symbol information of equalizers 39
2.2 Complexity of equalizer structures 54
2.3 Performance degradation due to channel estimation 56
2.4 Performance comparison of equalizer etructures for $\mathbf{c_1}$ 57

4.1 ξ vs. SNR for the LC-BiDFE tap optimization with the modified cost function for $\mathbf{c}_2$ 82

B.1 Training sequences for an EDGE system with $M = 2$ 130
B.2 Training sequences for an EDGE system with $M = 4$ 132

C.1 List of bird species seen in the Cayuga lake basin 134

List of Figures

1.1 Structure of a decision feedback equalizer 3
1.2 The equalization problem 7
1.3 Thesis organization: a road map 12

2.1 Baseband system model 15
2.2 Digital baseband equivalent of system 16
2.3 Structure of a GSM packet 16
2.4 Block diagram of a DFE 21
2.5 Block diagram of a DFE with ideal feedback assumption 21
2.6 Error propagation in DFE 28
2.7 Gap from the matched filter bound and ideal DFE 30
2.8 Structure of a Non-causal DFE 31
2.9 Structure of a Bidirectional DFE 37
2.10 Structure of a Bidirectional Arbitrated DFE 40
2.11 MSE performance for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$ 46
2.12 SER performance for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$ 47
2.13 SER performance comparison with MLSE for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$ 48
2.14 MSE performance curves for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$ 49
2.15 SER performance curves for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$ 50
2.16 Performance of a Bidirectional Arbitrated DFE (BAD) for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$ 51
2.17 Comparative MSE performance with same total number of taps for a BiDFE and a DFE 53
2.18 Effect of channel estimation on SER performance for an asymmetric channel with an impulse response of $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$ 55

3.1 Structure of an LC-BiDFE. 59

3.2 Performance gap from the matched filter bound for a symmetric 3-tap channel with root locations at $\left(\beta,\ \frac{1}{\beta}\right)$. 65
3.3 Performance gap from the matched filter bound for an asymmetric 3-tap channel with root locations at $\left(\beta,\ \frac{1}{2}\right)$. 66
3.4 LC-BiDFE structure for tap optimization in an infinite length scenario . 67

4.1 MSE performance of a finite length MMSE-BiDFE 79
4.2 SER performance of a finite length MMSE-BiDFE 80
4.3 SER performance of a finite length "modified" MMSE-BiDFE 82
4.4 Comparison of feedback filter tap weights 83
4.5 SER performance of an iterative finite length LC-BiDFE 85

5.1 BLAST scheme for a multi-element antenna system 90
5.2 Structure of a MIMO DFE 93
5.3 MSE performance of user 1 for the MIMO test channel C with $M = 2,\ P = 2$ 96
5.4 MSE performance of user 2 for the MIMO test channel C with $M = 2,\ P = 2$ 97
5.5 SER performance curves of user 1 for the MIMO test channel C with $M = 2,\ P = 2$ 98
5.6 SER performance curves of user 2 for the MIMO test channel C with $M = 2,\ P = 2$ 99

A.1 Power profile of fading channel for a typical urban environment 122
A.2 SER performance comparison for a fading channel environment 123
A.3 SER performance comparison for a few sample channels with the urban power profile 124
A.4 Channel zeros for $\mathbf{c}_3 = [-0.3920\ \ 0.0152\ \ 0.8320\ \ 0.0152\ \ -0.3920]$. 125
A.5 SER performance comparison for a deep null channel with impulse response $\mathbf{c}_3 = [-0.3920\ \ 0.0152\ \ 0.8320\ \ 0.0152\ \ -0.3920]$ 126
A.6 SER performance of an iterated LC-BiDFE for the channel with impulse response $\mathbf{c}_4 = [-0.2053\ \ 0.1300\ \ 0.9391\ \ 0.1300\ \ -0.2053]$ . . 127
A.7 SER performance comparison of users 1-4 for a MIMO fading channel environment with $M = 4,\ P = 4$ 128

B.1 Loss due to channel estimation, $M = 2$ 131
B.2 Loss due to channel estimation, $M = 4$ 133

Chapter 1

Introduction

In recent years, there has been an increase in demand on the data rate capabilities of wireless systems. This has necessitated an increase in bandwidth and signaling rate. As the bandwidth increases, the multipath distortion or frequency-selective fading caused by the physical medium becomes worse. The multipath channel causes a time dispersion of the transmitted signal, resulting in the overlap of the various transmitted symbols at the receiver. This is referred to as intersymbol interference (ISI), which, if left uncompensated, causes high error rates. A solution to the ISI problem is to design a receiver that employs a means for compensating or reducing the ISI in the received signal prior to detection. Such a compensator for the ISI is called an equalizer.

The delay spreads (in symbol durations) of wireless channels critically depends on the signaling bandwidth and physical environment. For instance, a GSM system, with a symbol rate of 270.83 Kbaud, in an urban environment has a typical delay spread of $5\mu s$. A digital terrestrial TV broadcast channel, on the other hand, uses a symbol rate of 10.78 Mbaud and can have worst case delay spreads in the order of $50\mu s$. The baseband digital impulse response for a DTV channel has a span of about 600 symbol taps and, hence, the equalizer is a very critical component of the receiver.

A significant number of wireless communication systems are packet based. In a GSM system, the information to be transmitted is encoded, interleaved, and then divided into small bursts, prior to transmission. At the receiver, each received signal burst is equalized and concatenated, prior to deinterleaving and decoding. The use of such *packet based communication systems* has made block processing of the received data possible. In this scenario, non-causal processing of the received signal becomes a possibility and in some cases may be advantageous.

1.1 Background

A powerful technique for combating ISI is Maximum-Likelihood Sequence Estimation (MLSE) [41], which minimizes probability of error events. This joint equalization and detection technique can be implemented with the trellis-based Viterbi algorithm [42]. However, the complexity of the Viterbi algorithm is proportional to the number of states in the trellis, which grows exponentially with the number of symbol intervals of the channel time dispersion. If the size of the symbol alphabet is $\mathcal{A}$ and the number of interfering symbols contributing to ISI is L, the Viterbi algorithm computes $\mathcal{A}^{L+1}$ branch metrics for each new received symbol. When the channel span is large, or if the alphabet size is big, then the MLSE approach becomes impractical. A number of schemes, e.g., [35, 57], generally known as reduced-state sequence estimation (RSSE), were proposed in the 1970s and 1980s in an attempt to approach the performance of MLSE at a reduced complexity.

One suboptimal scheme that is widely used in practice, due to its simple implementation, is the linear equalizer [61]. This approach employs a linear transversal filter and has a computational complexity that is (in a sense) a linear function of the channel dis-

persion length L. For a channel introducing only mild interference, the performance achievable by a conventional linear equalizer is often adequate. But, in suppressing the ISI, the linear equalizer inevitably enhances the noise. Therefore, as the channel distortion becomes severe such that there appear spectral nulls in the Nyquist band, the applicability of a linear equalizer is eventually limited by the noise enhancement or noise gain.

The basic limitation of a linear equalizer in coping with severe ISI has motivated a considerable amount of research into suboptimal nonlinear equalizers with low computational complexity, such as the *decision feedback equalizer* (DFE). The classic structure of this receiver is shown in Figure 1.1. The DFE consists of a linear feedforward filter (FFF) and a feedback filter (FBF). The FFF suppresses the contribution of the pre-cursor ISI, namely the interference caused by the symbols transmitted after the symbol of interest. The FBF cancels the post-cursor ISI by subtracting a weighted linear combination of the previous symbol decisions, assumed to be correct. The result is then applied to a threshold device to determine the symbol of interest. The FFF enhances the noise, but the noise gain is not as severe as in the case of a linear equalizer.

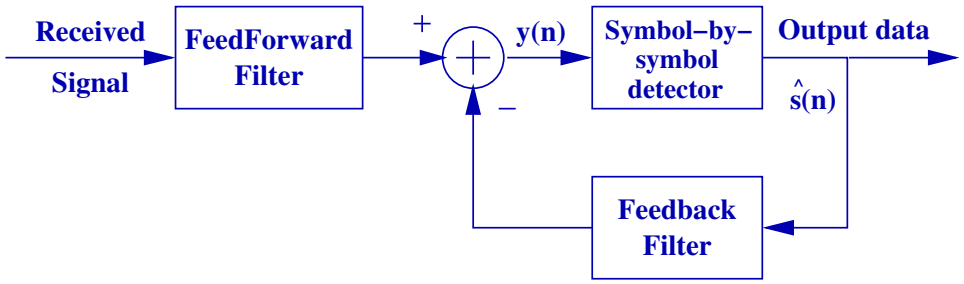

Figure 1.1: Structure of a decision feedback equalizer

The idea of using previous decisions to cope with the ISI problem was first introduced in 1967 by Austin [9], only two years after the development of the digital (adap-

tive) linear equalizer by Lucky [61]. The decision feedback receiver that minimizes the MSE between the input to the threshold device and the transmitted symbol was first obtained by Monsen [70]. In [74], the joint optimization of the transmitter and receiver, to maximize the output signal-to-noise ratio (SNR) of a zero-forcing DFE, was considered by Price. The jointly optimum MMSE transmitter and receiver was obtained by Salz [80]. Using linear space geometric arguments Messerschmitt [67] showed the equivalence of zero-forcing (ZF) decision feedback to MMSE prediction of a random process (and also the equivalence of a linear ZF receiver to linear interpolation of a random process), and thus provided simple derivations of the optimum filters and conditions for their existence. In [17], Belfiore and Park introduced a new DFE structure, called the noise predictive DFE and showed its equivalence to the conventional DFE for infinite-length filters. Infinite-length results on the MMSE-DFE were extended to the finite-length case in [1]. Thus, developments in DFE design have continued over 30 years since its inception.

The DFE design has typically been carried out assuming that the past decisions are error-free, thus simplifying the mathematics involved. However, when an error is made by the receiver the output of the FBF is no longer the desired value and the probability of subsequent errors is increased resulting in errors that tend to occur in bursts. The first residual-induced decision error, called a primary error, is fed back by the FBF causing secondary errors and creating an error burst. This phenomenon, known as *error propagation,* is more severe when the tap weights and/or the number of the feedback taps are large.

The performance loss of the DFE due to error propagation can be evaluated using Markov chain analysis. This technique, however, is not feasible when the FBF length is large as the number of states in the Markov chain increases exponentially with the FBF

length. Some research, which began in the early 1970s shortly after the invention of the DFE, has been done to simplify the analysis by clumping error states and providing upper bounds on the error probabilities. [4, 25, 53] use this approach to provide bounds on error propagation. [16, 52, 28, 73] are based on Markov chain techniques and provide other error measures such as mean error burst length. It is also possible to study error propagation from a dynamical systems perspective [54, 55].

Although the phenomenon of error propagation has been investigated for a long time, only a moderate amount of work has been done on mitigating error propagation in decision feedback equalizers. Several techniques, based on transmitter precoding [87, 46], were proposed in the early 1970s to remedy the error propagation problem in the DFE. These techniques require knowledge of the channel response at the transmitter, and thus are inapplicable if the transmission is broadcast or if the channel is time-varying. Dogncay [29] proposed techniques for detecting decision errors in equalization schemes including DFEs. In [36], Fertner introduced a scheme which attempts to prevent primary errors and hence error propagation. However, the performance of this algorithm critically depends on the value of the first post-cursor and the accompanying noise level. Ariyavisitakul proposed the use of a combination of soft decisions and delayed decoder decisions to cancel ISI for a joint convolutional coding and DFE scheme in [8]. In [8], the soft decision device for the DFE was obtained by using a simplified maximum a-posteriori probability (MAP) algorithm. In [10, 14] the use of a soft decision device to mitigate error propagation was considered and the decision device was optimized for minimizing the mean-squared error (MSE) and the bit-error rate (BER). In [31], a technique that combines a DFE with a Viterbi algorithm to provide a trade-off between complexity and performance was proposed.

Even in the presence of ideal feedback, i.e., no decision errors, the performance

of the DFE is suboptimal when compared to the matched filter bound. The matched filter bound (MFB) is defined as the output signal-to-noise ratio (SNR) when a matched filter is used at the receiver and only one symbol is transmitted (i.e., no ISI). This is equivalent to the performance that can be obtained through a channel where all the energy is concentrated in a single tap. The gap from the MFB is a good metric to evaluate MSE performance of equalizer structures. The performance gap from the matched filter bound is another drawback of the DFE.

The principle of non-causal decision feedback equalization (NCDFE) was first proposed by Proakis [75]. The NCDFE uses both past and future decisions to cancel all the ISI. The MMSE optimization of the NCDFE was derived by Gersho and Lim [43]. An estimate of the future transmitted symbols is obtained by the use of a preliminary equalizer and the introduction of an appropriate delay prior to the NCDFE. When no decision errors are made, the NCDFE attains the MFB and hence performs better than a DFE. However, the presence of decision errors can cause a performance degradation and only a fraction of the achievable performance improvement (when compared to the DFE) can be realized. The feedback filter of the NCDFE is effectively twice as long as compared to the DFE and results in increased error propagation.

A burst mode unbiased MMSE version of the NCDFE was proposed by Slock and de Carvalho in [81]. Later, they also proposed the use of soft decisions to decrease the effect of error propagation in [27]. In 2001, Chan and Wornell proposed a block-iterative DFE [21] which attempts to attain the MFB. The block iterative DFE is similar to the NCDFE, but unlike the NCDFE, the tap coefficients of the block iterative DFE are recomputed for each iteration to incorporate the effect of the reliability of the decisions that are used to cancel the pre-cursor and post-cursor ISI. The optimal DFE tap coefficients are derived under the assumption that the frequency response of the channel is i.i.d (independent and

identically distributed) and drawn from a complex Gaussian source. This assumption is satisfied only when the number of channel taps tends to infinity. Numerical results show that when the channel delay spreads are large (about 100 taps), the performance of the block-iterative DFE approaches the MFB after a few iterations.

1.2 Motivation

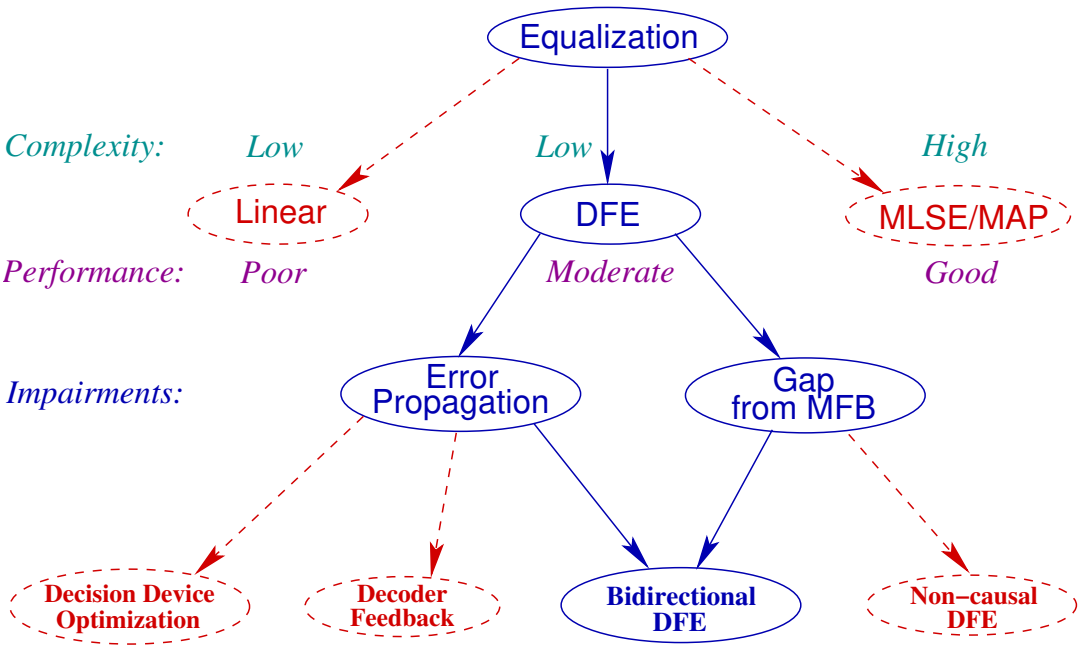

Figure 1.2: The equalization problem

The motivation for this dissertation has been pictorially represented in Figure 1.2. The low complexity and moderate performance of the DFE make it a preferred choice as an equalizer structure. However, the DFE has been shown to possess two significant drawbacks - error propagation and a gap from the matched filter bound. Some of the earlier work on DFEs have attempted to address these two issues on an individual basis. Recently, it has been shown [11, 64, 12] that the employment of time-reversal of the

received signal, along with bidirectional processing using a DFE structure, is an effective means of simultaneously addressing these two drawbacks and improving performance.

A time-reversal operation is done by reversing the sequential order of the received samples, in time, prior to equalization. The linear convolution structure imposed by the channel is conserved by the time-reversal operation. However, the equivalent channel impulse response, as seen by the equalizer, becomes a time-reverse of the actual channel impulse response. The idea of bidirectional equalization of the received signal employing a DFE was independently proposed in the early 90s by Ariyavisitakul [5] and Suzuki [83].

Ariyavisitakul [5, 6] proposed the use of a selective time-reversal structure in a DFE. He considered a packet transmission system communicating through a quasi-static frequency selective fading channel. The time-reversal structure was originally proposed to improve performance when a finite length constraint on the number of DFE filter taps is imposed. Under such a constraint, the performance of a normal mode DFE and a time-reversal mode DFE is different. Ariyavisitakul proposed the use of either a normal mode DFE or a time-reversal mode DFE, based on the performance of the two DFE structures over the channel realization encountered by the packet. It was further shown that if the length constraint on the DFE is relaxed, both the normal mode DFE and the time-reversal mode DFE have the same performance. A similar idea was also considered in [50] and [60]. Although the term bidirectional DFE is used in some of these publications, it refers to the selective time-reversal approach.

Recently, both Balakrishnan et al. [14] and McGahey et al. [64] have independently proposed a truly bidirectional DFE architecture. In a bidirectional DFE, henceforth referred to as BiDFE, the received signal is processed using both a normal mode and a time-reversal mode DFE and the output of the two streams are combined to improve

performance. In [64], the combining of the two output streams is done using a reconstruction based arbitration technique. The arbitration is performed to select, on a symbol-by-symbol basis, between the outputs of the normal mode and the time-reversal mode DFEs. As error propagation is a causal phenomenon, the error bursts of a normal mode DFE and a time-reversal mode DFE proceed in opposite directions in time. On the basis of numerical results, McGahey et al. reported that the improvement in their bidirectional arbitrated DFE (also known as BAD) is obtained by exploiting the low correlation between the error bursts caused by the normal mode DFE and the time-reversal mode DFE.

However, the advantage of combining the outputs of the normal mode DFE with that of the time-reversal mode DFE is not limited to mitigating error propagation. In fact, even in the presence of ideal feedback, namely no decision errors, the noise at the outputs of the normal mode DFE and the time-reversal mode DFE exhibit a low correlation. A diversity combining scheme employing a weighted linear combination of the outputs of the normal mode DFE and the time-reversal mode DFE was proposed in [11] and it was shown that the LC-BiDFE (linear combining bidirectional DFE) has a smaller value of noise-enhancement, when compared to either of the two constituent DFEs. In fact, if the finite length constraint is relaxed, the LC-BiDFE, under the ideal feedback assumption, actually attains the matched filter bound [12]. This "diversity" arises from the assumption that the past symbols are known to the normal mode DFE and the future transmitted symbols are known to the time-reversal mode DFE. The non-causal processing of the received signal along with the nonlinear structure of the DFE make this knowledge, although imperfect in the presence of decision errors, possible.

The idea of time-reversal has been employed in diverse areas in the past. Raphaeli [77] employs time-reversal to improve performance of a reduced state Viterbi algorithm

applied to demodulation of trellis codes. [78] extends this idea to reduced-state sequence estimators for combined equalization and decoding. Similar approaches employing time-reversal and a bidirectional structure [65, 58] have also been proposed for RSSEs. In [59] Lindskog employs time-reversal to extend the Alamouti scheme [3], that achieves transmit diversity for two transmit antennas over flat fading channels, to frequency-selective channels.

The demand for higher data rates in packet based communication systems necessitates improved signal processing at the receiver, albeit, at a low complexity. The purpose of this dissertation is to demonstrate that the *bidirectional decision feedback equalizer* (BiDFE) architecture proposed here is a suitable candidate for such an improved equalizer structure. The idea of using time-reversal in a DFE has received little attention from the research community and investigation of this topic would contribute towards promoting our understanding of the DFE and its unrealized potential.

The three BiDFE architectures, namely, the selective time-reversal DFE [5, 6], the LC-BiDFE [11, 12] and the BAD [64], each offer a performance/complexity trade-off that is distinct from that of the others. The primary focus of this dissertation will be the LC-BiDFE, the structure of which is particularly amenable to theoretical analysis; the intuition resulting from this study will help us better understand the other BiDFE architectures. The LC-BiDFE does not necessarily offer the best performance when compared to every other BiDFE architecture (e.g., BAD), but in most scenarios it provides a good performance at a significantly lower complexity.

1.3 Organization

A pictorial depiction of the organization[1] of this dissertation is presented in Figure 1.3. Chapter 2 describes the system model, notation and the various assumptions used in this dissertation. A brief review of the DFE is followed by a discussion on its limitations. Some of the conventional methods of addressing the drawbacks of the DFE are discussed. A generalized bidirectional DFE architecture that employs time-reversal of the received signal is proposed as an appealing alternative to simultaneously address the limitations of the DFE. The performance improvement offered by the BiDFE is demonstrated, both analytically and numerically. The LC-BiDFE is shown to offer significant performance improvement over the conventional DFE for only a modest increase in complexity. Various issues related to the implementation of the BiDFE are discussed. on this

The length constraint of an LC-BiDFE is relaxed and the performance for an infinite length scenario is analyzed in Chapter 3. The infinite length LC-BiDFE suffers from a smaller, but non-zero, gap from the MFB when compared with the MSE performance of an infinite length DFE. This motivates the tap optimization problem for the LC-BiDFE, where the coefficients of the constituent DFEs are optimized to minimize the overall MSE of the system, rather than the output MSE of the individual streams. The MMSE-BiDFE solution that minimizes the overall MSE of the LC-BiDFE is derived. It is shown that the infinite length MMSE-BiDFE, under the ideal feedback assumption, attains the matched filter bound.

In Chapter 4, the tap optimization problem is extended to the more practical scenario of a finite length LC-BiDFE. First, the tap coefficients of the LC-BiDFE are optimized

[1]The material in this dissertation is constituted from parts or whole of the papers [11, 12, 13, 89] written by the author in collaboration with others

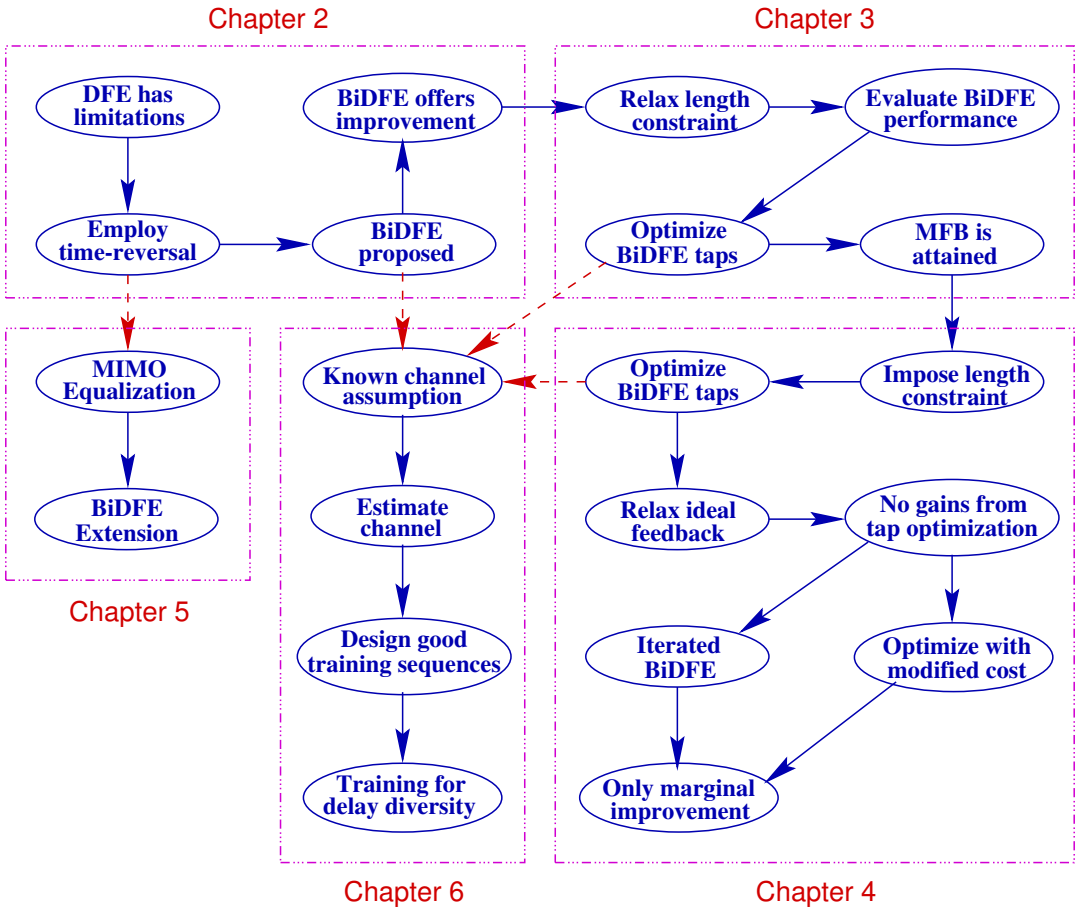

Figure 1.3: Thesis organization: a road map

for the special case of a symmetric channel. It is shown that the additional MSE performance improvement obtained via tap optimization, using the ideal feedback assumption, does not readily translate into an improvement in symbol error rate (SER) performance when decision feedback is employed. Two solutions, one based on tap optimization with a modified cost and another on an iterative approach, are suggested to overcome this drawback in the absence of ideal feedback. The BiDFE tap optimization is also extended to the asymmetric channel case.

The idea of employing time-reversal is extended in Chapter 5 to the multiple-input multiple-output (MIMO) channel equalization problem. The MIMO DFE affords the

possibility of user re-ordering in addition to time-reversal. A MIMO BiDFE structure that incorporates both time-reversal and user re-ordering is proposed and is shown to provide performance improvements similar to the single-input single-output (SISO) BiDFE.

Chapters 2-5 assume that the channel impulse response is perfectly known at the receiver. In practice this assumption is almost always violated and the channel response is typically estimated with the aid of a training sequence. The design of training sequences for the SISO channel estimation problem has been closely studied for over three decades. However, very little is known on the design of optimal training sequences for the MIMO channel estimation problem and this is considered in Chapter 6. The optimality criterion for the MIMO training sequences is derived, and the design trade-offs in the choice of training sequence length are discussed. A few reduced complexity techniques, for the implementation of a search for good training sequences, are proposed and near-optimal training sequences determined using these methods are tabulated. For the special case of a delay-diversity scheme, it is shown that transmitting the same optimal training sequence from each of the two transmit antennas is optimal for training based channel estimation.

Finally, Chapter 7 provides a summary of results and directions for future research. MATLAB script files for reproducing the simulation plots in this dissertation are available at "http://bard.ece.cornell.edu/matlab/balakrishnan/index.html".

Chapter 2

Bidirectional DFE

This chapter reviews some well known results on DFEs and a few earlier attempts to address its drawbacks. Section 2.1 describes the system model and states the assumptions made in this dissertation. The matched filter bound (MFB), a suitable metric to evaluate mean-squared error (MSE) performance of equalizer structures, is defined in this section. Section 2.2 introduces the structure of a DFE, tap optimization of the filter taps, the gap from the matched filter bound, and the error propagation phenomenon observed in DFEs. The non-causal DFE structure proposed by Gersho and Lim [43], to attain the matched filter bound, is described in Section 2.3. A few techniques that mitigate error propagation in DFEs are also reviewed. The concept and implication of time-reversal of the received signal is introduced in Section 2.4. The bidirectional DFE structure is proposed in Section 2.5 as a suitable solution to address the limitations of a DFE. The rationale behind the BiDFE architecture, the BAD architecture proposed by McGahey et al. in [64] and the LC-BiDFE proposed in [11], is discussed. The performance improvement offered by the BiDFE is investigated, both analytically and numerically, and various implementation issues are discussed. Section 2.6 provides a summary of the contents presented in this chapter.

2.1 System Model

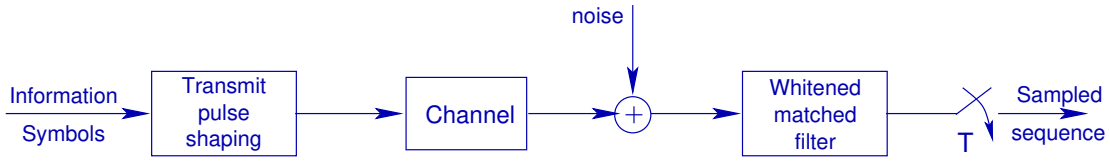

Figure 2.1: Baseband system model

Data communication through a frequency-selective channel is considered. Typically, a digital communication system consists of an information source, a source encoder and a channel encoder. The output of the channel encoder is mapped onto a signal constellation. The mapped symbols are then pulse shaped, modulated by a carrier, and transmitted through a frequency-selective channel. At the receiver, the received signal is down-converted, filtered (with an appropriate receive filter), sampled and processed. This system can be modeled using the baseband equivalent depicted in Figure 2.1. In the baseband equivalent system shown in Figure 2.1, the information symbols are transmitted through a frequency-selective, baseband, analog channel and are corrupted by additive noise at the receiver. A whitened matched filter front-end is employed at the receiver and the output of the matched filter is sampled at baud-rate (symbol rate) to obtain a sampled sequence. Forney [41] has shown that the sampled sequence at the receiver constitutes a sufficient statistic for detecting the information symbols. Hence, any optimal receiver structure can be preceded by the whitened matched filter [41]. Furthermore, the presence of a whitening front-end filter ensures that the noise component of the sampled sequence is white.

The digital baseband equivalent of the analog channel model is illustrated in Figure 2.2. The channel C represents the digital equivalent of the combined transmit pulse-shaping filter, analog channel and front-end whitened matched filter. However, the chan-

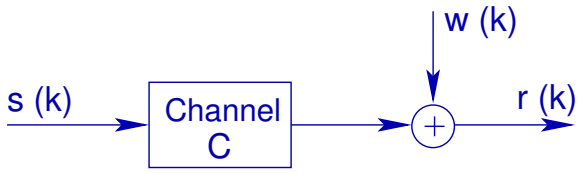

Figure 2.2: Digital baseband equivalent of system

nel impulse response is not known a priori at the receiver and is usually estimated using a training sequence embedded in the data. Hence, it is not feasible to have a front-end analog filter that is matched to the channel impulse response. Typically, a front-end receive filter that is matched to the transmit pulse shaping function is used and the output of the receive front-end filter is oversampled to obtain sufficient statistics. For example, when a square-root-raised-cosine (SRRC) function is used as a pulse shaping filter this same filter is the typical front-end architecture of the receiver and it has the added advantage of aiding timing synchronization. Yet, we assume the digital baseband equivalent model shown in Figure 2.2; the intuition and results derived with such a model are general enough and can be easily extrapolated to accommodate oversampling at the receiver.

Figure 2.3: Structure of a GSM packet

We consider a packet based communication system model, where N is the number of symbols transmitted in each data packet. The structure of a typical packet for a Global Systems Mobile (GSM) transmission standard is illustrated in Figure 2.3. A

GSM packet has a length of about 156 symbols with the encoded data split into two chunks of 58 symbols each. A 26 symbol long training sequence is embedded in the center of the packet as a mid-amble and is used for channel estimation. The remainder of the packet consists of tail bits and guard symbols. The packetization of the transmitted data results in some edge effects, i.e., the symbols at the edge experience a truncated version of the channel impulse response as opposed to the symbols at the center of the packet. Kaleh [49] and de Carvalho [81, 27] have discussed the implication of the edge effects in packet based systems and shown that this can be exploited to decrease the symbol-error-rate (SER) for the symbols at the edges (or) the symbols adjoining the training sequences. However, as this aspect of the packet based system is limited to the edge symbols, it will not be discussed here. Unless stated otherwise, we make the following assumptions.

Assumption 2.1 *The channel C can be modeled as a linear time-invariant (LTI) filter for the duration of each packet.*

Most wireless communication channels are time-varying, albeit slowly. Although, the time variations can be significant between non-contiguous packets, the channel variations over the duration of a packet can be ignored under certain conditions. For instance, in a wireless environment under low mobility scenarios, this assumption is valid and is referred to as a quasi-static fading channel [6]. In a wireline environment, the time variations are typically slow, and if the duration of the packet is small, compared to the time-constant of channel variations, this assumption is valid.

Assumption 2.2 *The channel can be modeled as a finite impulse response (FIR) filter.*

This is a common assumption used in the literature. Although, the delay spreads of the channel impulse response depends on the transmission environment, the FIR approximation is typically valid. For a GSM transmission system, the channel delay spread can

be as much as 20 μs (equivalent to 20 symbol periods for a data rate of 1 Mbps) for a hilly terrain.

For the causal linear channel model with a finite span, i.e., finite-impulse-response (FIR), the received sequence $r(k)$ can be expressed as,

$$r(n) = \sum_{k=0}^{L_c-1} c(k)s(n-k) + w(n) \tag{2.1}$$

where $\mathbf{c} = [c(0)\ c(1)\ \ldots\ c(L_c-1)]^T$ is the channel impulse response with L_c taps, $s(n)$ represents the transmitted source symbols and $w(n)$ is the additive noise sequence. From equation (2.1), we notice that at any sampling instant the received sequence is a function of not only the current transmitted symbol, but also contains components from the earlier transmitted symbols. This phenomenon is known as intersymbol interference (ISI).

Assumption 2.3 *The noise sequence $w(n)$ is additive white Gaussian (AWGN) with a variance σ_w^2 and is independent from the source sequence $s(n)$.*

This is a standard assumption made in the literature. The digitization of the received signal results in the truncation of the probability distribution function (pdf) of the noise process. Although, the pdf of the noise process has only a finite support, it will be approximately Gaussian.

Assumption 2.4 *The channel impulse response* **c** *is known at the receiver.*

In typical communication systems, the channel impulse response is not known at the receiver. Most packet based communication systems provide a training segment, which could be in the order of 10-20% of the packet size, in the data burst. A priori knowledge of the training sequence is used at the receiver to estimate the channel impulse response. The accuracy of the channel estimate is dependent on the length of the training sequence

relative to the delay spread of the channel impulse response, and the noise variance. However, we assume perfect knowledge of the channel at the receiver.

2.1.1 Matched Filter Bound

The matched filter receiver was originally designed as an optimal detector for messages transmitted using a set of orthonormal basis functions [76, 47]. Turin [88] provides a review of the matched filter and its properties. In the context of digital communication through a frequency-selective channel, the sampled output of a matched filter receiver provides sufficient statistics for detection. Here, the sufficient statistic bears all the information (about the transmitted symbols) contained in the received signal. For the system illustrated in Figure 2.2, the matched filter consists of a digital FIR front-end with an impulse response given by $\tilde{\mathbf{c}}^*$, where $\tilde{c}(k) = c(-k)$. Let $y(n)$ denote the output of such a matched filter receiver. Then $y(n)$ can be expressed as,

$$\begin{aligned} y(n) &= \sum_{k=-L_c+1}^{0} c^*(-k) r(n-k) \\ &= \sum_{k=-L_c+1}^{L_c-1} \phi(k) s(n-k) + \sum_{k=-L_c+1}^{0} c^*(-k) w(n-k) \end{aligned} \tag{2.2}$$

where

$$\phi \triangleq \mathbf{c} * \tilde{\mathbf{c}}^* \tag{2.3}$$

is the auto-correlation function of the digital channel impulse response. When only one symbol is transmitted in a packet, i.e., $N = 1$, the interfering symbols are all zero. Hence, there is no contribution from the ISI terms in equation (2.2), and it reduces to,

$$y(n) = \phi(0) s(n-k) + \sum_{k=-L_c+1}^{0} c^*(-k) w(n-k) \ . \tag{2.4}$$

In such a scenario, the matched filter receiver maximizes the signal-to-noise ratio (SNR) at the output. The output SNR of such a matched filter receiver is commonly referred to

as the matched filter bound (MFB) and is given by,

$$\gamma_{\text{mfb}} = \frac{\sigma_s^2 \phi(0)}{\sigma_w^2} \tag{2.5}$$

where σ_s^2 is the variance of the source symbols. The MFB can be attained at the receiver under the following scenarios,

1. When only one symbol is transmitted in the packet. This results in no intersymbol interference and the matched filter receiver is the optimal detector.

2. If all transmitted symbols in the packet, except the symbol of interest, are known. In this case, the ISI components affecting the symbol of interest at the output of the matched filter receiver (see equation (2.2)) can be canceled perfectly.

In reality, when the above two conditions are not satisfied, the matched filter bound cannot be attained under most circumstances. Even the optimal equalizer, namely the MLSE, has a gap from the MFB in most scenarios and this gap can be determined based on the coefficients of the channel impulse response [76]. In fact, the MLSE attains the MFB only when the channel impulse response is limited to a length of 2 symbol taps. In spite of the fact that the MFB cannot be attained, the gap from the MFB is a useful metric to compare equalizer structures.

2.2 DFE Review

The decision feedback equalizer consists of a feedforward filter (FFF), a feedback filter (FBF) and a decision device. The DFE block diagram is illustrated in Figure 2.4. Let the feedforward filter F have an impulse response, $\mathbf{f} = [f(0)\ f(1)\ \ldots\ f(L_f - 1)]^T$, where L_f is the number of taps in the FFF. Similarly, the feedback filter B is an L_b-tap filter with an impulse response, $\mathbf{b} = [b(1)\ b(2)\ \ldots\ b(L_b)]^T$. The decision device, $\mathcal{Q}$ is

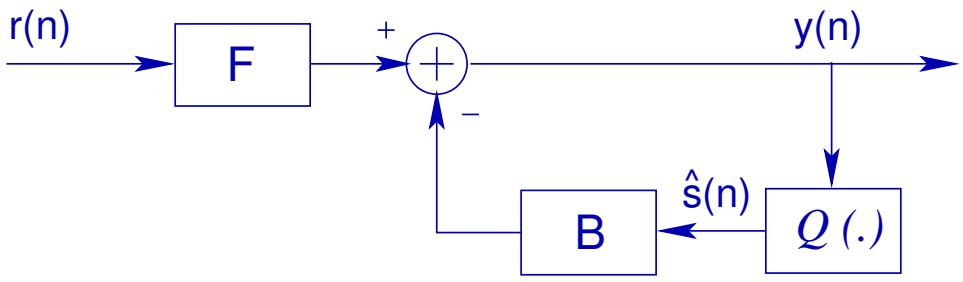

Figure 2.4: Block diagram of a DFE

usually a quantizer that makes hard decisions $\hat{s}(n)$ on the output $y(n)$ of the DFE. The output $y(n)$ of the DFE can then be mathematically expressed as,

$$y(n) = \sum_{k=0}^{L_f-1} f(k)r(n-k) - \sum_{k=1}^{L_b} b(k)\hat{s}(n-k) \ . \tag{2.6}$$

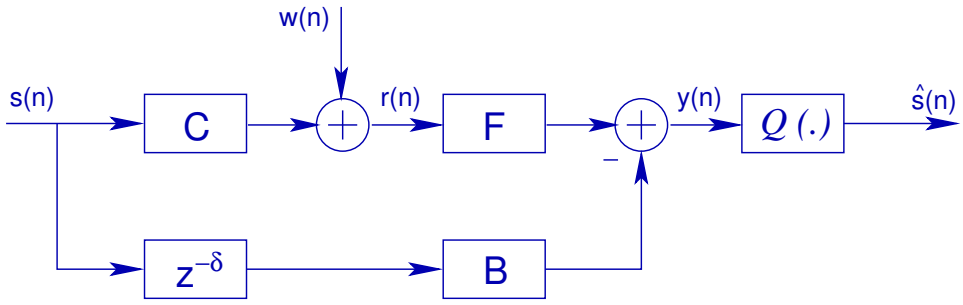

Figure 2.5: Block diagram of a DFE with ideal feedback assumption

Assumption 2.5 *The decisions fed back to the FBF are error-free, i.e., ideal decision feedback.*

Although, this is almost never true in a practical implementation, it is a key assumption made while designing DFEs. The "ideal feedback" assumption simplifies the mathematics involved and makes analysis tractable. However, the effect on the symbol error rate performance, when this assumption is violated, will also be investigated, albeit numerically.

Let δ be the detection delay of the DFE. Then, under assumption 2.5, the equivalent block diagram of the DFE is shown in Figure 2.5 and equation (2.6) can be rewritten as

$$y(n) = \sum_{k=0}^{L_f-1} f(k)r(n-k) - \sum_{k=1}^{L_b} b(k)s(n-\delta-k). \tag{2.7}$$

Let $\mathbf{h}$ represent the combined channel-feedforward filter impulse response, i.e, $\mathbf{h} = \mathbf{c} * \mathbf{f}$. Then, $\mathbf{h} = [h(0)\ h(1)\ \ldots\ h(L_c+L_f-2)]^T$ has a length $L_h \triangleq L_c+L_f-1$ taps. Equation (2.7) can be rewritten as,

$$\begin{aligned} y(n) = \underbrace{h(\delta)s(n-\delta)}_{\text{cursor}} &+ \underbrace{\sum_{k=0}^{\delta-1} h(k)s(n-k)}_{\text{res. pre-cursor ISI}} + \underbrace{\sum_{k=1}^{L_b}\{h(k+\delta)-b(k)\}s(n-\delta-k)}_{\text{modeled post-cursor ISI}} \\ &+ \underbrace{\sum_{k=\delta+L_b+1}^{L_c+L_f-2} h(k)s(n-k)}_{\text{residual post-cursor ISI}} + \underbrace{\sum_{k=0}^{L_f-1} f(k)w(n-k)}_{\text{filtered noise}} \end{aligned} \tag{2.8}$$

An effective way of equalizing the channel would be for the FFF to

- shape the combined channel-feedforward filter response $\mathbf{h}$ to maximize the energy of the cursor, i.e., $h(\delta) \approx 1$ in the unbiased design
- result in a small residual ISI, i.e., $h(k) \approx 0$ for $0 \leq k < \delta$ and $\delta + L_b < k \leq L_c + L_f - 2$
- keep the noise gain, $\mathbf{f}^T\mathbf{f}$ as small as possible

and for the FBF to

- cancel exactly the modeled post-cursor ISI by matching the FBF taps to that of the combined response, i.e., $b(k) = h(k+\delta), 1 \leq k \leq L_b$.

This is essentially what the MSE optimized DFE does, as will become apparent in the next subsection.

2.2.1 Finite length MMSE-DFE

The coefficients of the DFE can be designed to be either zero-forcing (perfectly cancel the ISI) or to minimize the mean-squared error (MSE). The MMSE optimized DFE is popular and the finite length design was derived by Al-Dhahir [1]. A simplified derivation of the finite length optimal DFE coefficients will be reviewed in this subsection. The MSE at the output of the DFE represented in Figure 2.5 is given by,

$$\sigma_\eta^2 \triangleq \mathrm{E}\left[\{s(n-\delta) - y(n)\}^2\right] \ , \tag{2.9}$$

where $\eta(n) = s(n-\delta) - y(n)$ is the soft decision error of the equalizer. The goal of the MMSE design of the DFE is to determine the filter pair $(\mathbf{f}^{\text{mmse}}, \mathbf{b}^{\text{mmse}})$ that minimizes σ_η^2 of equation (2.9). We now define the $L_h \times L_f$ channel convolution matrix C as

$$C \triangleq \begin{bmatrix} c(0) & 0 & \cdots & 0 \\ \vdots & c(0) & \cdots & \vdots \\ c(L_c-1) & \vdots & \ddots & 0 \\ 0 & c(L_c-1) & \ddots & c(0) \\ \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & \cdots & c(L_c-1) \end{bmatrix} . \tag{2.10}$$

We also define the vectors $\mathbf{s}_n$, $\hat{\mathbf{s}}_n$ and $\mathbf{w}_n$ as,

$$\begin{aligned} \mathbf{s}_n &\triangleq [s(n)\ s(n-1)\ \ldots\ s(n-L_h-1)]^T \\ \hat{\mathbf{s}}_n &\triangleq [s(n-\delta-1)\ \ldots\ s(n-\delta-2)\ s(n-\delta-L_b)]^T \\ \mathbf{w}_n &\triangleq [w(n)\ w(n-1)\ \ldots\ w(n-L_f-1)]^T. \end{aligned} \tag{2.11}$$

From equations (2.1) and (2.6), equation (2.9) now becomes,

$$\sigma_\eta^2 = \mathrm{E}\left[\{s(n-\delta) - \mathbf{s}_n^T C\mathbf{f} + \hat{\mathbf{s}}_n^T\mathbf{b} - \mathbf{w}_n^T\mathbf{f}\}^2\right] \tag{2.12}$$

which can be minimized by setting the gradient of σ_η^2 with respect to $\mathbf{f}$ and $\mathbf{b}$ as zero. First,

$$\nabla_{\mathbf{b}}\sigma_\eta^2 = \mathrm{E}\left[-2\hat{\mathbf{s}}_n\mathbf{s}_n^T C\mathbf{f} + 2\hat{\mathbf{s}}_n\hat{\mathbf{s}}_n^T\mathbf{b} - 2\hat{\mathbf{s}}_n\mathbf{w}_n^T C\mathbf{f} + 2s(n-\delta)\hat{\mathbf{s}}_n\right]. \tag{2.13}$$

Assumption 2.6 *The source symbols are independent and identically distributed (i.i.d) with a variance* σ_s^2.

From assumptions 2.3, 2.5 and 2.6, equation (2.13) simplifies to,

$$\nabla_{\mathbf{b}}\sigma_\eta^2 = -2\sigma_s^2 MC\mathbf{f} + 2\sigma_s^2\mathbf{b} \tag{2.14}$$

where

$$M = [0_{L_b\times\delta}\ I_{L_b\times L_b}\ 0_{L_b\times L_b - L_h - \delta}]. \tag{2.15}$$

Hence, the MMSE FBF filter taps are

$$\mathbf{b}^{\text{mmse}} = MC\mathbf{f}. \tag{2.16}$$

Similarly, for the FFF parameters we have

$$\nabla_{\mathbf{f}}\sigma_\eta^2 = 2\sigma_s^2 C^T C\mathbf{f} - 2\sigma_s^2 C^T M^T\mathbf{b} - 2\sigma_s^2 C^T\mathbf{e}_\delta + 2\sigma_w^2\mathbf{f} \tag{2.17}$$

where $\mathbf{e}_\delta = [0\ \ldots\ 0\ 1\ 0\ \ldots\ 0]^T$ is an L_h length unit vector with a unity at position $\delta+1$, with $0 \le \delta < L_h$. Substituting the result of equation (2.16) into equation (2.17) yields

$$\nabla_{\mathbf{f}}\sigma_\eta^2 = 2\sigma_s^2 C^T(I - M^T M)C\mathbf{f} - 2\sigma_s^2 C^T\mathbf{e}_\delta + 2\sigma_w^2\mathbf{f}. \tag{2.18}$$

The MMSE feedforward equalizer is

$$\mathbf{f}^{\text{mmse}} = (C^T VC + \lambda I)^{-1} C^T\mathbf{e}_\delta \tag{2.19}$$

where

$$\lambda = \sigma_w^2/\sigma_s^2, \quad V = (I - M^T M). \tag{2.20}$$

AlDhahir [1] emphasizes the importance of an unbiased DFE, in which the cursor $h(\delta) = 1$. Although the MSE of an unbiased DFE design is larger than the MMSE-DFE, it has a lower probability of error. The unbiased MMSE-DFE can be obtained by minimizing the MSE of the DFE under a unit tap constraint for the cursor. It can be shown that, the unbiased MMSE-DFE taps are obtained from the biased MMSE-DFE by normalizing the tap coefficients by the cursor term, i.e.,

$$\mathbf{f}_u^{\text{mmse}} = \frac{\mathbf{f}^{\text{mmse}}}{\left(C\mathbf{f}^{\text{mmse}}\right)^T \mathbf{e}_\delta}, \ \mathbf{b}_u^{\text{mmse}} = \frac{\mathbf{b}^{\text{mmse}}}{\left(C\mathbf{f}^{\text{mmse}}\right)^T \mathbf{e}_\delta}. \tag{2.21}$$

2.2.2 Infinite length MMSE-DFE

In subsection 2.2.1 the MMSE-DFE solution was derived under a finite length constraint on the feedforward and feedback filters of the DFE. If the length constraint on the DFE filters is relaxed, we obtain the infinite length or the ideal DFE solution. In [69] Monsen formulated the optimum infinite length DFE under the MMSE criterion. Let $C(z)$ denote the Z-transform of the channel impulse response, namely

$$C(z) \triangleq \sum_{k=0}^{L_c-1} c(k) z^{-k}. \tag{2.22}$$

Let $F(z)$ and $B(z)$ denote the FFF and FBF in Z-transform representation. Then, under the ideal feedback assumption 2.5, the optimum filters $F(z)$ and $B(z)$ are given by,

$$F(z) = \frac{C^*(z^{-1})}{U^-(z)U^-(0)} \tag{2.23}$$

and

$$B(z) = \frac{U^+(z)}{U^-(0)} - 1 \tag{2.24}$$

where

$$U(z) = C(z)C^*(z^{-1}) + \lambda. \tag{2.25}$$

The terms $U^+(z)$ and $U^-(z)$ correspond to the minimum phase (all roots inside the unit circle) and the maximum phase (all roots outside the unit circle) spectral factors of $U(z)$, respectively. The term $U^-(0)$, in equations (2.23) and (2.24), refers to the constant term in the polynomial expansion of the maximum phase spectral factor $U^-(z)$. The feedforward filter can only be implemented as a stable, anti-causal filter with an infinite impulse response, while the FBF is strictly causal. From equation (2.23), it is clear that the FFF is composed of a filter matched to the channel impulse response and a whitening filter. When the noise variance is very small, i.e., $\sigma_w^2 \to 0$, then the FFF approaches an all-pass filter that reflects the maximum phase roots of the channel impulse response across the unit circle into minimum phase roots.

2.2.3 Gap from the MFB

In subsection 2.1.1, it was mentioned that the MFB can be attained only when all the interfering symbols are perfectly known at the receiver. The DFE, under the ideal feedback assumption, has perfect knowledge of at least the interfering symbols transmitted prior to the symbol being detected. Hence, one can expect the infinite length DFE to perform better or at least as well as an infinitely long linear equalizer. This is indeed true and has been shown in [69, 70]. However, the DFE still has a gap from the MFB and this gap was quantified for an infinite-length DFE by Ariyavisitakul in [6]. This will be discussed in greater detail in Chapter 3. To demonstrate the gap from the MFB, we consider the following example.

Example 2.1 *Consider a real, 2-tap channel with an impulse response* $C(z) = 1+az^{-1}$, *with the second tap of magnitude less than unity, i.e.,* $|a| < 1$. *Let the input SNR of the system be high, such that* $\lambda \triangleq \sigma_w^2/\sigma_s^2 \to 0$. *Then, the infinite length MMSE-DFE*

coefficients are given by,

$$F(z) \approx 1,\ B(z) \approx az^{-1} \tag{2.26}$$

and the output SNR of the DFE is

$$\gamma_{dfe} = \frac{\sigma_s^2}{\sigma_w^2}. \tag{2.27}$$

However, the MFB defined in equation (2.5) for this channel is

$$\gamma_{mfb} = \frac{\sigma_s^2(1+a^2)}{\sigma_w^2}. \tag{2.28}$$

The infinite length MMSE-DFE has an output SNR that is lower by a factor $(1+a^2)$ *from the MFB, for this channel example.*

2.2.4 Error Propagation

As long as the decisions made by the decision device are correct, the performance of the DFE is as desired (though short of the MFB). However, in the presence of noise and residual ISI, decision errors are inevitable. The first noise and residual ISI induced error is known as a primary error. As the decision error is fed back through the FBF, instead of canceling the post-cursor ISI components, it adds additional interference. This results in an increased probability of error during the subsequent symbol decisions. Hence, a primary error typically induces a burst of errors. The burst terminates when the residual errors add up in the right way [52, 54] to flush out all the decision errors from the FBF tapped delay line. The average duration of the error burst is a function of the channel and DFE coefficients, the source constellation and the noise level. Figure 2.6 illustrates the burstiness of the error propagation phenomenon in a DFE. The source symbols are drawn from a binary phase-shift-keying (BPSK) constellation. The soft output of the DFE is plotted in Figure 2.6 along with a "*" to mark the locations of the hard decision

errors. The degradation due to error propagation generally increases with the magnitude of the FBF coefficients and the length of the FBF.

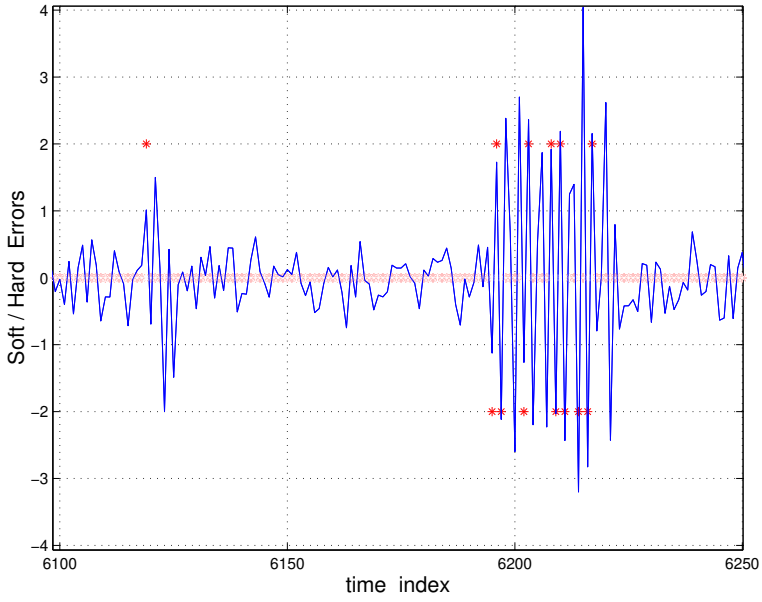

Figure 2.6: Error propagation in DFE

2.2.5 Numerical Example

The gap from the matched filter bound and the degradation due to error propagation for a DFE has been illustrated in Figure 2.7. The symbol error rate (SER) curves have been plotted as a function of the SNR in Figure 2.7. The source symbols were drawn from a BPSK constellation and transmitted through a channel with an impulse response $\mathbf{c} = [0.2887 \quad 0.9129 \quad 0.2887]$. The DFE was allocated 4 taps for the feedforward filter and 3 taps for the feedback filter. The performance curves for an ideal DFE and a conventional DFE are plotted along with the matched filter bound. For the ideal DFE, perfect decisions, i.e., error-free decisions, are fed back to the FBF. To plot the SER curve corresponding to the matched filter bound, assumption 2.3 was invoked. The

symbol error rate for the matched filter bound is given by,

$$\text{SER}_{\text{mfb}} = Q\left(\frac{1}{\sigma_w}\right) , \tag{2.29}$$

where σ_w^2 is the noise variance and $Q(\cdot)$ is the Q-function defined as,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(\frac{-z^2}{2}\right) dz. \tag{2.30}$$

The performance degradation of the conventional DFE with respect to the matched filter bound can be decomposed into two parts. The first, namely *error propagation gap* is the performance degradation of the conventional DFE when compared to the ideal DFE. For an SER of 10^{-3}, the error propagation gap is about 0.7 dB. The second component of the DFE performance degradation is known as the *gap from the matched filter bound* and is given by the performance difference of the ideal DFE curve and the MFB. In this example, the gap from MFB is about 1.5 dB at an SER of 10^{-3}. Note that the gap from the matched filter bound increases with increasing SNR. This is due to the fact that at high SNR scenarios, ISI is the dominant factor.

2.3 DFE Enhancements

There have been a few attempts in the literature to address the drawbacks of the DFE, namely the gap from the MFB and error propagation, individually. A few of these attempts were briefly discussed in Chapter 1. In this section, some of the enhancements proposed to the DFE structure to mitigate these problems are discussed in greater detail.

2.3.1 Non-causal DFE

As indicated in subsection 2.1.1, the MFB can be attained when all interfering symbols are perfectly known. In 1970, Proakis [75] proposed the principle of non-causal decision

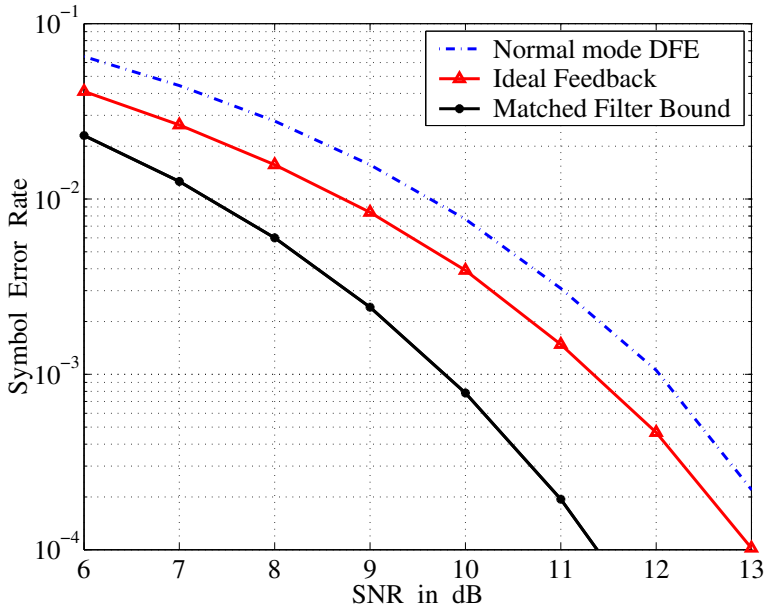

Figure 2.7: Gap from the matched filter bound and ideal DFE

feedback equalization (NCDFE) to attain the MFB. The MMSE optimized NCDFE taps were derived by Gersho and Lim [43]. The block diagram of an NCDFE is depicted in Figure 2.8. The NCDFE consists of a preliminary equalizer in parallel with the DFE structure. The preliminary equalizer is used to provide estimates of the "future symbols" to the NCDFE. The preliminary equalizer can be either a linear equalizer or even a DFE. The DFE component of the NCDFE is preceded by a delay element and consists of a feedforward filter F and a feedback filter B. Typically, the delay preceding the DFE is chosen to be at least as large as the detection delay of the preliminary equalizer. Then, the interference due to the "future symbols", i.e., the pre-cursor ISI, at the output of the filter F can be canceled using the estimates $\hat{s}_1(n)$ from the preliminary equalizer. The decisions $\hat{s}(n)$ of the NCDFE are used to cancel the post-cursor ISI component.

Gersho and Lim have shown in [43] that for the MSE optimized NCDFE, the forward filter F is proportional to the matched filter $C^*(z^{-1})$ and the feedback filter B is a

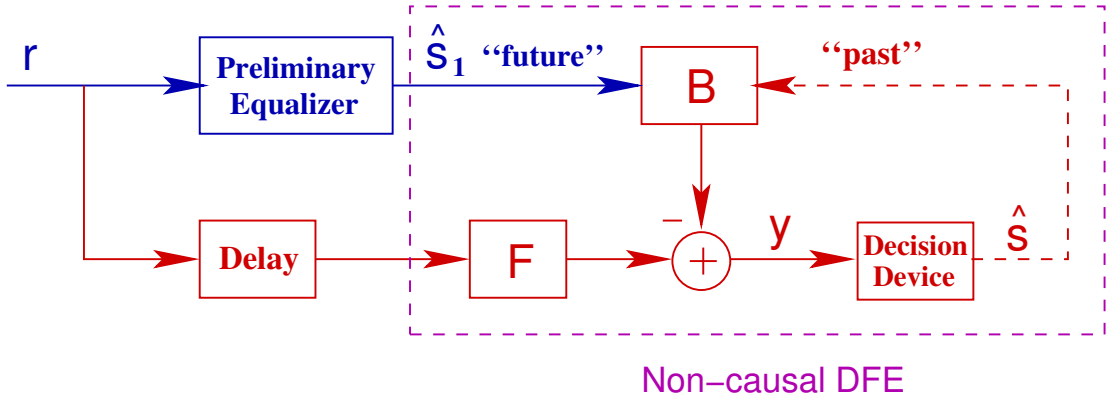

Figure 2.8: Structure of a Non-causal DFE

cascade of the channel and feedforward filter impulse response, without the central coefficient. In the presence of decision errors, the NCDFE suffers from an error propagation phenomenon that results in performance degradation. Unlike the DFE, the NCDFE cancels both the pre-cursor and post-cursor ISI using the symbol decisions. Hence, the FBF in an NCDFE is, in effect, twice as long as that of a DFE. As the degradation due to error propagation is more severe with a longer FBF, some of the performance improvement expected due to closing the gap from the MFB is lost to increased error propagation gap. In [27], Slock and de Carvalho have proposed the use of soft decisions and an iterative NCDFE approach to further improve the performance of the NCDFE. In the iterative approach, the more reliable symbol decisions, i.e., $\hat{s}(n)$, of the NCDFE are used as a tentative estimate of the "future symbols" to cancel the pre-cursor ISI component for another NCDFE. This operation is performed iteratively, in an attempt to refine the symbol decisions. Although, these solutions are useful in decreasing the effect of error propagation, they come at the cost of increased complexity.

In [20, 21], Chan and Wornell have proposed a block-iterative DFE, which attempts to attain the MFB. The block iterative DFE has a structure very similar to the NCDFE, but instead of using decisions from a preliminary equalizer, the iterative DFE uses deci-

sions from the previous iteration. Furthermore, the tap coefficients of the block iterative DFE are recomputed during each iteration to incorporate the effect of the reliability of the decisions that are used to cancel the pre-cursor and post-cursor ISI. This work, however, assumes that the channel impulse response is infinitely long, i.e., $L_c \rightarrow \infty$ and the frequency response of the channel is drawn from a complex Gaussian random process. Numerical results show that when the channel delay spreads are large (over 100 taps), the high SNR performance of the block-iterative DFE, after a small number of iterations, tends towards the MFB. It is not clear whether these results readily translate for practical channels (say, for delay spreads typical to wireless channels).

2.3.2 Decoder Feedback

In systems employing error control codes, a popular idea to address the error propagation problem in DFEs has been to use feedback from the decoder. As the decisions at the output of the channel decoder are more reliable than the quantizer decisions, the decoder decisions can be fed back through the FBF to mitigate error propagation. However, there is a delay associated with the channel decoder and hence Ariyavisitakul [8, 7] proposed the idea of splitting the FBF of the DFE into two parallel filters, one of which uses the decoder decisions to cancel the post-cursor ISI components corresponding to symbols $[s(n-D-\delta),\ s(n-D-\delta-1),\ \ldots]$, where D is the decoder delay, while the other cancels the remainder of the post-cursor ISI using the tentative DFE decisions.

An alternate method of decoder feedback proposed in [10, 62] is the iterative DFE-decoder architecture, referred to as the *Turbo DFE*. In the Turbo DFE, the received signal is equalized using a DFE and then the transmitted bits are decoded using a channel decoder. The more reliable decoder decisions are now used as the FBF inputs for another DFE that operates on a delayed version of the input signal. The equalization-decoding

operation can be performed iteratively to mitigate error propagation. Numerical results, provided in [10], show that the error propagation gap can be decreased by about 75% with a small number of iterations.

2.3.3 Decision Device Optimization

The replacement of the decision device, shown as a quantizer in the DFE block diagram of Figure 2.4, with a soft decision device is an effective technique to mitigate error propagation for both coded and uncoded systems. The use of a soft decision device has been proposed in [8, 27, 10, 14]. In [10, 14], Balakrishnan et al. optimize the soft decision device in the presence of error propagation under both the minimum MSE and the minimum bit-error rate (BER) criteria. For instance, the soft decision device (see [14] for details) that minimizes MSE at the DFE output, in the presence of error propagation, for a BPSK source constellation is a sigmoid function given by $\mathcal{Q}(y) = \tanh(y/\sigma_\eta^2)$, where σ_η^2 is the variance of the filtered noise and residual ISI.

2.4 Time-reversal

Consider the operation of time-reversal applied to the received sequence. A time-reversal operation on the sequence $r(n)$ of equation (2.1) can be mathematically expressed as,

$$\begin{aligned} \tilde{r}(n) &\triangleq r(-n) \\ &= \sum_{k=-(L_c-1)}^{0} \tilde{c}(k)\tilde{s}(n-k) + \tilde{w}(n) \end{aligned} \tag{2.31}$$

where "$\sim$" implies time-reversal. Now,

$$\tilde{c}(k) \triangleq c(-k), \quad \tilde{s}(n) \triangleq s(-n), \quad \text{and} \quad \tilde{w}(n) \triangleq w(-n) \tag{2.32}$$

are the time-reversed channel impulse response, source sequence and noise sequence, respectively. In equation (2.31), the linear convolution structure of the channel is preserved. Hence, this is equivalent to transmitting the time-reversed source sequence through a channel impulse response, whose coefficients are the time-reversed version of $\mathbf{c}$.

Property 2.1 *The time-reversed channel has root locations that are reciprocal of the roots of the original channel impulse response, i.e., the roots are reflected across the unit circle.*

Proof: Recall the z-transform representation of equation (2.22). We let $\beta_i, \ \forall \ i = 1, \ \ldots, \ L_c - 1$ represent the zeros of the FIR channel $C(z)$. Then,

$$C(z) = \beta_0 \prod_{i=1}^{L_c-1} (1 - \beta_i z^{-1}) \tag{2.33}$$

where $\beta_0 = c(0)$ is the first tap coefficient of the channel. Let $\tilde{C}(z)$ be the z-transform of the time-reversed channel impulse response $\tilde{\mathbf{c}}$. Then $\tilde{C}(z) = C(z^{-1})$ and can be expressed as,

$$\begin{aligned} \tilde{C}(z) &= \beta_0 \prod_{i=1}^{L_c-1} (1 - \beta_i z) \\ &= \beta_0 \prod_{i=1}^{L_c-1} (-\beta_i z) \left(1 - \frac{z^{-1}}{\beta_i}\right). \end{aligned} \tag{2.34}$$

Hence, $1/\beta_i, \ \forall \ i = 1, \ \ldots, \ L_c - 1$ are now the zeros of the time-reversed channel impulse response. ■

Effectively, the time-reversal operation converts the minimum phase roots of the channel in to maximum phase roots and vice-versa.

2.4.1 Selective Time-Reversal DFE

Property (2.1) was exploited by Ariyavisitakul in [5, 6] to improve the performance of a DFE with a finite length constraint, in a packet based system. Ariyavisitakul proposed the use of two parallel DFE structures, one of which is a normal mode DFE that operates on the received signal. The received packet of data is time-reversed and another DFE is used to equalize the time-reversed received sequence. Ariyavisitakul argued that under a finite length constraint, one of the two DFE structures would have a better performance. If the channel impulse response is known at the receiver (see assumption 2.4) then the finite length MMSE-DFE coefficients can be determined and the performance of the normal mode and time-reversal mode DFEs determined. Based on the MSE performance metric, Ariyavisitakul chooses the output of one of the DFE structures for the duration of the packet. It was further shown that, as the number of DFE taps increases, the difference in performance between the normal mode and time-reversal mode DFE structures diminish and the advantage of using the *selective time-reversal structure* for the DFE disappears. The difference in performances of the two DFE modes can be illustrated by the following example.

Example 2.2 *Consider a real, 2-tap channel with an impulse response* $C(z) = a + z^{-1}$*, with the first tap of magnitude less than unity, i.e.,* $|a| < 1$*. Let us constrain the FFF and FBF taps to be of length one each. Then the unbiased MMSE-DFE taps are,*

$$F(z) = \frac{1}{a},\ B(z) = \frac{z^{-1}}{a} \tag{2.35}$$

and the output SNR of the normal mode DFE is

$$\gamma_{dfe} = \frac{\sigma_s^2}{a^2\sigma_w^2}. \tag{2.36}$$

On the other hand, for the time-reversed channel impulse response $\tilde{C}(z)$*, the unbiased*

MMSE-DFE taps are

$$F_{tr}(z) = 1,\ B_{tr}(z) = az^{-1} \tag{2.37}$$

and the output SNR of the time-reversal mode DFE is

$$\gamma_{trdfe} = \frac{\sigma_s^2}{\sigma_w^2}. \tag{2.38}$$

Since, $|a| < 1$, the output SNR of the time-reversal mode DFE is greater than that of the normal mode DFE and the selective time-reversal structure will choose the outputs of the time-reversal mode DFE.

2.5 Bidirectional DFE

The idea of time-reversal in a DFE can be extended from selective time-reversal to a bidirectional DFE (BiDFE). The block diagram of a BiDFE is similar to that of a selective time-reversal DFE, but with an additional diversity combining block that combines the outputs of the normal mode and time-reversal mode DFE. The structure of a BiDFE, proposed in [11], is illustrated in Figure 2.9. The BiDFE processes the received packet of data in two parallel streams. The received sequence $r(n)$ is equalized in stream I using a DFE with an inherent detection delay δ_1. The feedforward filter $\mathbf{f_1}$ has L_{f1} taps, while the feedback filter $\mathbf{b_1}$ has L_{b1} taps. The combined channel-feedforward filter impulse response for the normal mode DFE, namely stream I, is denoted by $\mathbf{h_1} = \mathbf{c} * \mathbf{f_1}$.

A block time-reversal operation is applied to the received sequence $r(n)$ in stream II. The only difference in the block time-reversal operation, when compared to the time-reversal operation defined in equation (2.32), is a time-shift operation that manifests as an addition of a constant term to the index of the time-reversed sequence. We now define,

$$\tilde{r}(n) \triangleq r(N + L_c - n) \quad \forall n = 1,\ \ldots,\ N + L_c - 1 \tag{2.39}$$

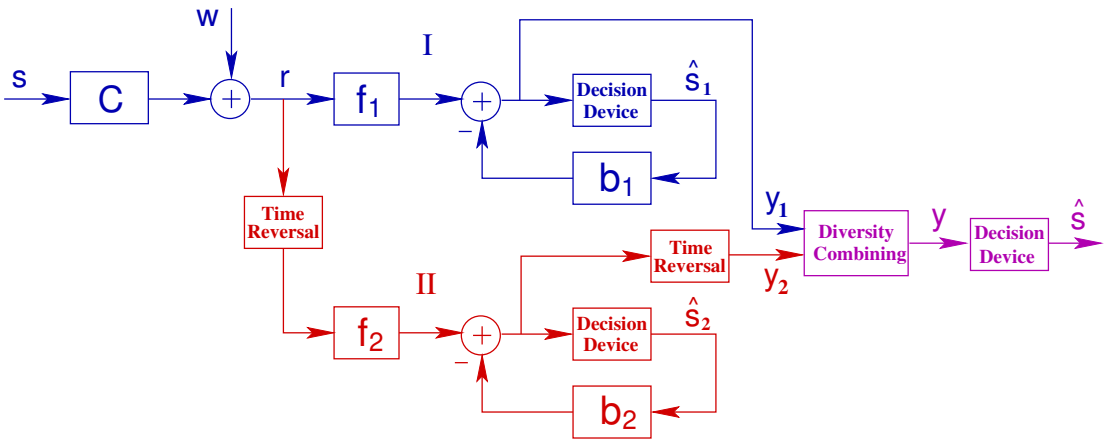

Figure 2.9: Structure of a Bidirectional DFE

From equation (2.1), equation (2.39) can be rewritten as

$$\tilde{r}(n) = \sum_{k=0}^{L_c-1} c(k)s(n + L_c - n - k) + w(N + L_c - n) \tag{2.40}$$

Now, we define the block time-reversed source sequence, channel response and noise sequence as

$$\tilde{s}(n) \triangleq s(N - n + 1), \quad \forall\, n = 1,\ \ldots,\ N, \tag{2.41}$$

$$\tilde{c}(k) \triangleq c(L_c - 1 - k), \quad \forall\, k = 0,\ \ldots,\ L_c - 1, \tag{2.42}$$

$$\tilde{w}(n) \triangleq w(N + L_c - n), \quad \forall\, n = 1,\ \ldots,\ N + L_c - 1, \tag{2.43}$$

respectively. With a few simple manipulations, equation (2.40) can be simplified to

$$\tilde{r}(n) = \sum_{k=0}^{L_c-1} \tilde{c}(k)\tilde{s}(n - k) + \tilde{w}(n) \tag{2.44}$$

where "$\sim$" implies block time-reversal. Let $\mathbf{f_2}$ with L_{f2} taps and $\mathbf{b_2}$ with L_{b2} taps be the feedforward and feedback filter settings of the time-reversal mode DFE, of stream II, that equalizes the time-reversed channel $\tilde{\mathbf{c}}$. We further define the combined channel-feedforward filter impulse response for the time-reversal mode DFE as $\mathbf{h_2} = \tilde{\mathbf{c}} * \mathbf{f_2}$ with δ_2 denoting the detection delay.

2.5.1 Rationale

What is the rationale in going from the selective time-reversal DFE of subsection 2.4.1 to the BiDFE structure proposed in Figure 2.9? The advantages of the BiDFE structure are two-fold. Firstly, error propagation is a *causal* phenomenon. A primary error induces a burst of secondary errors that proceed in the forward direction in time. When the signal is processed using a time-reversal mode DFE, the error propagation runs in backward time. Based on the fact that the error bursts run in opposite directions in time for the normal mode and the time-reversal mode DFE, one can expect a low correlation between these error bursts. Perhaps, this feature of the BiDFE can be exploited to mitigate error propagation.

Recall from subsection 2.1.1 that the MFB can be attained when all the interfering symbols, namely the past and the future interfering symbols are perfectly known. This is the reason why an NCDFE attains the MFB. Under the ideal feedback assumption 2.5, the proposed BiDFE structure has perfect knowledge of both the past symbols (due to the normal-mode DFE) and the future symbols (due to the time-reversal mode DFE). Hence, one can expect the BiDFE to have a lesser gap from the MFB than a conventional DFE. The known symbol information for various equalizer structures has been compared in Table 2.1.

2.5.2 BAD: Bidirectional Arbitrated DFE

The bidirectional arbitrated DFE (BAD) proposed by McGahey et al. in [64] is a special case of the BiDFE with an elaborate diversity combining block. The diversity combining block is shown in Figure 2.10 and consists of a reconstruction and arbitration stage. In the reconstruction stage, the symbol decisions $\hat{s}_1(n)$ and $\hat{s}_2(n)$ of the normal mode and time-reversal mode DFEs are filtered through an estimate of the channel impulse

Table 2.1: Known symbol information of equalizers

Equalizer Structure	**Past symbols**	**Future symbols**
Linear Equalizer	Unknown	Unknown
DFE	Known	Unknown
BiDFE	Known	Known
NCDFE	Known	Known

response to reconstruct the received sequence. The reconstructed received sequences, namely $\hat{r}_1(n)$ and $\hat{r}_2(n)$ are compared to the true received sequence $r(n)$ and a windowed distance metric is computed. For instance,

$$\hat{d}_1(n) = \sum_{k=-K}^{k=K} |\hat{r}_1(n-k) - r(n-k)|^2. \tag{2.45}$$

The windowed distances $\hat{d}_1(n)$ and $\hat{d}_2(n)$ are compared and the output of the DFE resulting in a smaller windowed distance is selected as the output of the arbiter for that symbol instant, i.e.,

$$\hat{s}(n) = \begin{cases} \hat{s}_1(n), & \text{if} \quad \hat{d}_1(n) < \hat{d}_2(n) \\ \hat{s}_2(n), & \text{if} \quad \hat{d}_1(n) > \hat{d}_2(n) \end{cases} . \tag{2.46}$$

BAD was proposed to exploit the directionality of error propagation to improve performance and numerical results demonstrating performance improvements were provided in [64].

2.5.3 Preliminaries

In this subsection, the mathematical model that leads to the derivation of a simple linear diversity combing scheme is provided. For the BiDFE, we make the following simplifying assumptions.

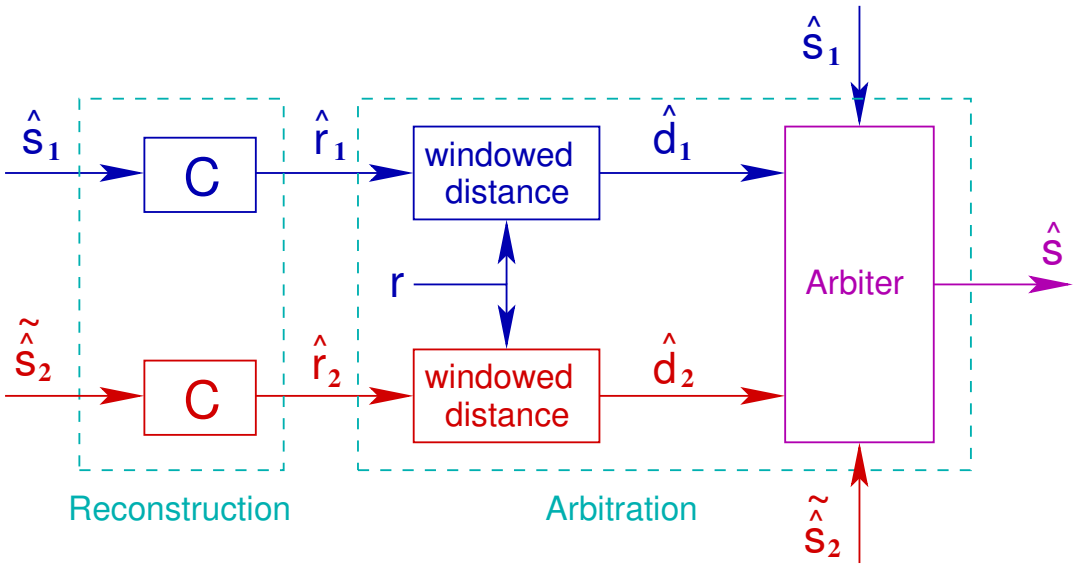

Figure 2.10: Structure of a Bidirectional Arbitrated DFE

Assumption 2.7 *The DFEs for the normal mode and time-reversal mode streams are finite length unbiased MMSE-DFEs, i.e.,* $h_1(\delta_1) = h_2(\delta_2) = 1$.

As the output SNR, or equivalently the gap from the MFB, will be used as the metric to evaluate equalizer structures, it is meaningful to assume the DFEs to be unbiased and also to minimize the MSE.

Assumption 2.8 *The source sequence* $s(n)$ *has unit variance,* $\sigma_s^2 = 1$.

This assumption can be made without loss of generality as our interest is in the SNR term and not on any absolute values.

Let $y_1(n)$ be the soft output (also the input to the decision device) of the normal mode DFE and $y_2(n)$ be the block time-reversed (and synchronized) soft output of the time-reversal mode DFE. The sequences $y_1(n)$ and $y_2(n)$ can be expressed as,

$$y_1(n) = s(n) + \eta_1(n) \tag{2.47}$$

$$y_2(n) = s(n) + \eta_2(n) \tag{2.48}$$

where, under the ideal feedback assumption 2.5

$$\eta_1(n) = \sum_{m=0}^{L_{f1}-1} f_1(m)w(n+\delta_1-m) + \sum_{m=0}^{\delta_1-1} h_1(m)s(n+\delta_1-m) + \sum_{m=\delta_1+L_{b1}+1}^{L_{f1}+L_c-2} h_1(m)s(n+\delta_1-m) \tag{2.49}$$

$$\eta_2(n) = \sum_{m=0}^{L_{f2}-1} f_2(m)w(n+L_c-1-\delta_2+m) + \sum_{m=0}^{\delta_2-1} h_2(m)s(n-\delta_2+m) + \sum_{m=\delta_2+L_{b2}+1}^{L_{f2}+L_c-2} h_2(m)s(n-\delta_2+m). \tag{2.50}$$

Let σ_1^2 and σ_2^2 be the variances of $\eta_1(\cdot)$ and $\eta_2(\cdot)$, respectively. We use ρ to denote the coefficient of correlation between the sequences $\eta_1(\cdot)$ and $\eta_2(\cdot)$, i.e.,

$$\rho = \frac{\mathrm{E}[\eta_1\eta_2]}{\sigma_1\sigma_2} \tag{2.51}$$

where E[·] is the expectation operation.

2.5.4 Diversity Combining

In subsection 2.5.2, we discussed an elaborate reconstruction based arbitration technique for the diversity combining block. Although, this technique could be quite effective in exploiting the low correlation between the error bursts, it is extremely hard to analyze. Furthermore, it is not clear what fraction of the performance gain of the BAD is due to error propagation mitigation and how much is due to a decreased gap from the MFB. We now consider a much simplified and low complexity diversity combing block which generates the outputs $y(n)$ as a weighted linear combination of the sequences $y_1(n)$ and $y_2(n)$, namely

$$\begin{aligned} y(n) &= \alpha y_1(n) + (1-\alpha)y_2(n) \\ &= s(n) + \alpha\eta_1(n) + (1-\alpha)\eta_2(n) \end{aligned} \tag{2.52}$$

where α is the weighting factor. The architecture employing this weighted linear combination of the outputs of the constituent DFEs will be referred to as a linear-combining bidirectional DFE (LC-BiDFE). Although, such a simplified diversity combiner is ill-equipped to handle the bursty nature of error propagation, we can obtain useful intuition on how well the gap from the MFB has been decreased. The MSE for the soft output $y(n)$ is

$$\begin{aligned} \text{MSE} &= \text{E}[|y(n) - s(n)|^2] \\ &= \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 + 2\rho\alpha(1-\alpha)\sigma_1\sigma_2 \ . \end{aligned} \tag{2.53}$$

Lemma 2.1 *The weighting factor α, that minimizes the MSE of equation (2.53) is given by*

$$\alpha^{opt} = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \tag{2.54}$$

Proof: To minimize the MSE of equation (2.53), set the first derivative of the MSE with respect to α to zero. Now,

$$\frac{\partial \text{MSE}}{\partial \alpha} = 2\alpha\sigma_1^2 - 2(1-\alpha)\sigma_2^2 + 2\rho(1-2\alpha)\sigma_1\sigma_2. \tag{2.55}$$

Setting the gradient of MSE to zero, i.e.,

$$\frac{\partial \text{MSE}}{\partial \alpha} = 0, \tag{2.56}$$

we obtain the optimal weighting factor of equation (2.54). When $\sigma_1 = \sigma_2$, the optimal weighting factor is $\alpha = \frac{1}{2}$ and this is known as equal-gain combining. ■

Furthermore, it can be shown that the minimum MSE obtained with the optimal value of α is

$$\text{MMSE} = \frac{(1-\rho^2)\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \tag{2.57}$$

The MMSE of equation (2.57) is less than the MSE of both the normal mode and the time-reversal mode DFEs. Hence, the performance of a LC-BiDFE is better than the performance of a DFE with selective time-reversal structure.

2.5.5 Time-Reversal Diversity

In Lemma 2.1, it was demonstrated that the LC-BiDFE has a better MSE performance than each of the constituent DFEs. In this subsection, a symmetric channel impulse response will be considered and the LC-BiDFE will be shown to result in a smaller value of noise gain and residual ISI terms. For a symmetric channel, the time-reversed channel impulse response $\tilde{\mathbf{c}}$ is equal to the channel impulse response $\mathbf{c}$. Hence, the same DFE tap settings can be used for the two streams, i.e., $\mathbf{f_1} = \mathbf{f_2} = \mathbf{f}$ and $\mathbf{b_1} = \mathbf{b_2} = \mathbf{b}$. Further, the mean-squared error (MSE) for the two streams, σ_1 and σ_2, are equal and the optimal weighting factor of equation (2.54) is $\alpha = \frac{1}{2}$. The soft outputs, $y(n)$ of the diversity combining block can be expressed as,

$$y(n) = s(n) + \eta(n) \tag{2.58}$$

where $\eta(n)$ is the sum of the noise and the residual ISI components and,

$$\eta(n) = \frac{\eta_1(n) + \eta_2(n)}{2}. \tag{2.59}$$

We define the vector $\mathbf{g}$,

$$g(k) = \begin{cases} 0, & \delta \leq k \leq \delta + L_b \\ h(k), & \text{otherwise} \end{cases} . \tag{2.60}$$

Equation (2.59) can be rewritten as,

$$\begin{aligned} \eta(n) = \frac{1}{2} \sum_{k=0}^{L_h-1} g(k)\{s(n+\delta-k) + s(n-\delta+k)\} \\ + \frac{1}{2} \sum_{k=0}^{L_f-1} f(k)\{w(n+\delta-k) + w(n+L_c-1-\delta+k)\} . \end{aligned} \tag{2.61}$$

The variance (or equivalently the MSE) of the overall system error of equation (2.61) can be expressed as a sum of the noise gain and the residual ISI terms.

$$\text{Noise gain term} = \sum_{k=k_{min}}^{k_{max}} \left\{ \frac{f(k) + f(2\delta + 1 - L_c - k)}{2} \right\}^2 \sigma_w^2 \tag{2.62}$$

where

$$k_{min} = \min(0, 2\delta + 1 - L_h) \quad \text{and} \quad k_{max} = \max(L_f - 1, 2\delta + 1 - L_c) \tag{2.63}$$

and

$$\text{Residual ISI term} = \sum_{k=j_{min}}^{j_{max}} \left\{ \frac{g(k) + g(2\delta - k)}{2} \right\}^2 \tag{2.64}$$

where

$$j_{min} = \min(0, 2\delta + 1 - L_h) \quad \text{and} \quad j_{max} = \max(L_h - 1, 2\delta) \ . \tag{2.65}$$

For each of the individual streams, the noise gain term and the residual ISI terms are $\sigma_w^2(\mathbf{f}^T\mathbf{f})$ and $\mathbf{g}^T\mathbf{g}$, respectively. From equations (2.62) and (2.64), we observe that the equal-gain diversity combining scheme has decreased[1] the contributions of both the noise gain and the residual ISI terms on the effective MSE when compared to either the normal-mode or time-reversal mode operation of the DFE. This results in a decreased gap from the MFB for the LC-BiDFE structure.

2.5.6 Simulation Results

In this subsection, the MSE and SER performance improvements provided by the BiDFE architecture (LC-BiDFE and BAD) over the conventional DFE structure will be investigated numerically, with the aid of two sample channel impulse responses. Additional simulation examples are provided in Appendix A. Matlab script files for generating the

[1]This can be easily proved using the Cauchy-Schwartz inequality.

simulation plots provided in this section are available at [19]. The SER performance of the LC-BiDFE will also be compared with BAD and the NCDFE. We first consider a real, asymmetric channel with an impulse response of

$$\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]. \tag{2.66}$$

Each of the constituent DFEs in the BiDFE structure were restricted to 4 feedforward taps and 3 feedback taps. As stated in assumption 2.7, unbiased MMSE-DFEs were used. The source symbols were selected from a BPSK source constellation. The MSE performance curves for a normal mode DFE, time-reversal mode DFE and the LC-BiDFE are compared in Figure 2.11. Correct decision feedback (assumption 2.5) was invoked for computing the MSE values. In Figure 2.11, the normal mode DFE has a marginally better performance than the time-reversal mode DFE. Ariyavisitakul's DFE with selective time-reversal will choose the normal mode DFE for this example. The LC-BiDFE, on the other hand, has a significantly better performance than either of the two DFE modes and yields a gain of about 1 dB for an MSE of 0.1 (i.e., -10 dB).

The symbol error rate (SER) curves for the normal mode DFE, time-reversal mode DFE and the LC-BiDFE are shown in Figure 2.12 for the same asymmetric channel impulse response $\mathbf{c}_1$. These simulation results do not invoke assumption 2.5, and hence, incorporate the effect of error propagation, i.e., decision feedback was assumed. In Figure 2.12, there are two performance curves shown for the LC-BiDFE, namely with a hard/soft DFE. In a soft DFE, the hard-limiter or quantizer of a hard DFE is replaced by a scheme with a soft decision device [14], which for a BPSK source constellation is a $\tanh(\cdot)$ nonlinearity. As can be seen from the plot, the LC-BiDFE structure incorporating a soft decision device has a gain of about 1.2 dB at an SER value of 10^{-3}. Although, the design of the diversity combining block of the LC-BiDFE structure does not take the error propagation into account, the performance gain observed in the presence of ideal

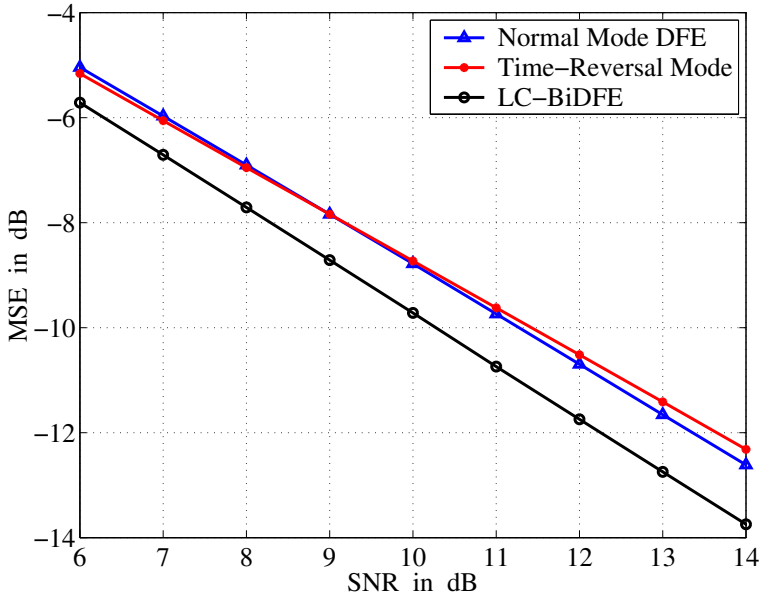

Figure 2.11: MSE performance for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$

feedback also translates to the case with decision feedback. Also, as expected from the discussion in subsection 2.3.3, the performance of a LC-BiDFE with soft decision feedback is better than that of a LC-BiDFE with hard decision feedback.

In Figure 2.13, the SER performance curves of the normal mode DFE and the LC-BiDFE (with soft decisions) are compared with the BAD and MLSE. A window of size $K = 5$, was used in BAD to compute the averaged Euclidean distance metric (see equation 2.45) in the arbitration stage. The MLSE was implemented using the Viterbi algorithm. Compared to the 1.2 dB of improvement offered by the LC-BiDFE (with soft decisions), over the normal mode DFE at an SER of 10^{-3}, the BAD offers an additional improvement of only 0.2 dB. Furthermore, the gap from the MLSE for the LC-BiDFE is 0.5 dB. It should be noted that the MLSE is the optimal detector for the transmitted symbols.

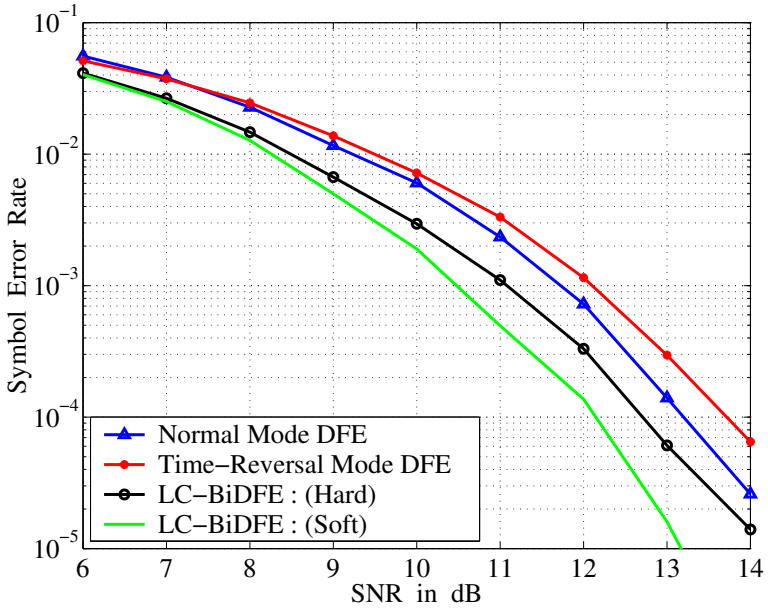

Figure 2.12: SER performance for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$

In another experiment a symmetric baud-spaced channel with an impulse response of

$$\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887] \tag{2.67}$$

was considered. A finite length constraint of 4 feedforward taps and 3 feedback taps was imposed on the constituent DFEs of the BiDFE structure. As $\mathbf{c_2}$ is a symmetric channel, the DFE settings, and hence the performance, are the same for both the normal mode and the time-reversal mode DFEs. Figure 2.14 compares the MSE performance of the DFE and LC-BiDFE with the MFB. Correct decision feedback (assumption 2.5) was invoked for computing the MSE values. The LC-BiDFE has a MSE performance gain of about 1 dB over the conventional DFE structure at an MSE of 0.1, but still suffers from a gap of 0.6 dB from the MFB.

The SER performance of a conventional DFE with and without ideal feedback is

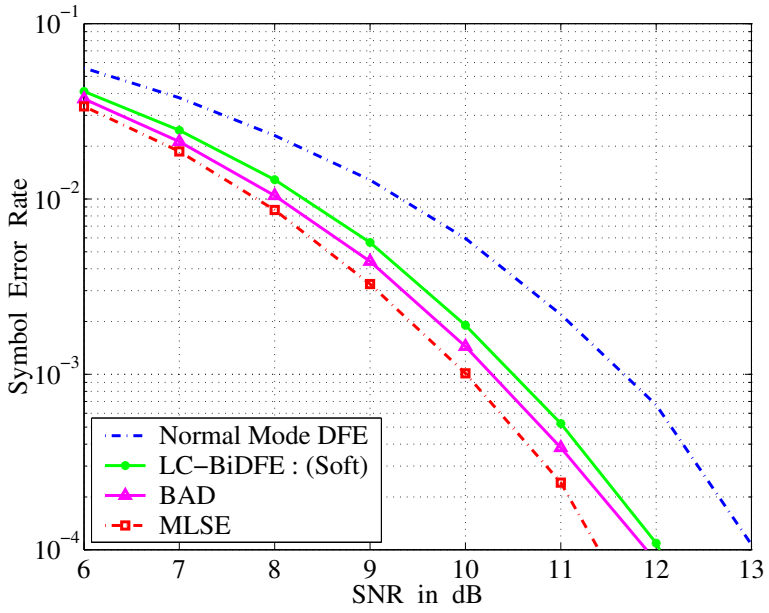

Figure 2.13: SER performance comparison with MLSE for an asymmetric channel with an impulse response $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$

compared with the performance of a LC-BiDFE and an NCDFE, both with decision feedback, in Figure 2.15. The error propagation gap for this example is nearly 0.7 dB and at an SER value of 10^{-3}, the LC-BiDFE (with hard decisions) has a performance gain of 0.7 dB over a conventional DFE. The NCDFE, on the other hand, has only a marginal performance improvement of 0.25 dB over the conventional DFE. This is in tune with the increased error propagation in an NCDFE that was discussed in subsection 2.3.1. For the NCDFE, a conventional DFE with 4 FFF taps and 3 FBF taps was used as a preliminary equalizer. The optimal feedforward filter for the NCDFE is a channel matched filter. Hence, 3 taps were allocated to the FFF of the NCDFE and 4 taps to the FBF (to cancel all the pre-cursor and post-cursor ISI terms) of the NCDFE.

As the bidirectional arbitrated DFE (BAD) has a diversity combining scheme more suited towards exploiting the burstiness of error propagation, in addition to the property

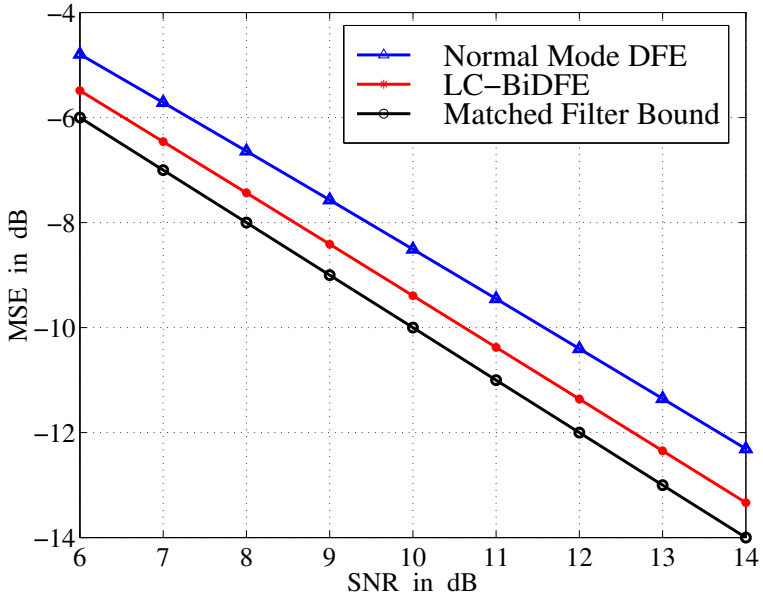

Figure 2.14: MSE performance curves for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$

of closing the gap from the MFB that is inherent to the BiDFE structure, the SER performance of a BAD is compared with the performance of a LC-BiDFE (with and without soft decisions), the normal mode DFE, and the MLSE in Figure 2.16. As expected the BAD outperforms the soft LC-BiDFE by about 0.3 dB at an SER of 10^{-3}, while the soft LC-BiDFE has a performance gain of 0.9 dB over the normal mode DFE performance. The MLSE, on the other hand, offers a performance improvement of 0.5 dB over the LC-BiDFE (with soft decisions). The performance improvement of BAD comes at the cost of additional complexity, due to the elaborate diversity combining structure. Figure 2.16 illustrates that near-optimal performance can be attained at a significantly lower complexity by employing the Bidirectional DFE architecture (BAD or LC-BiDFE).

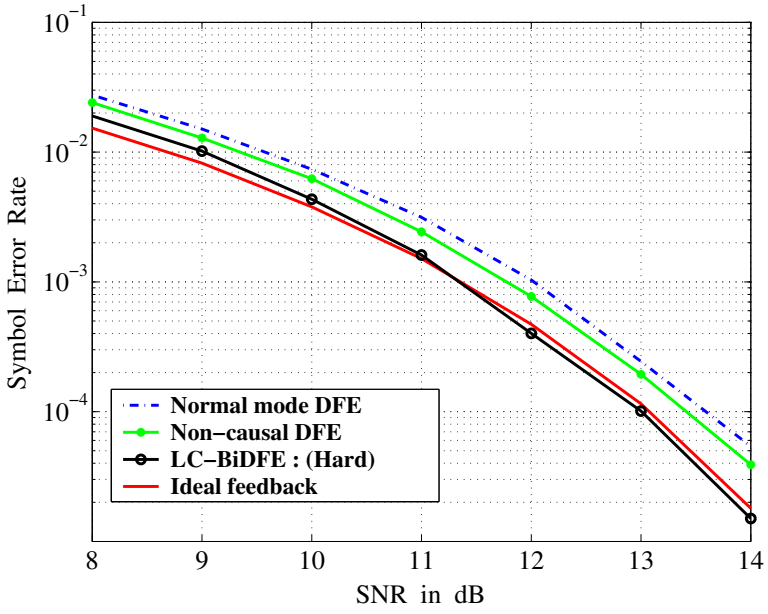

Figure 2.15: SER performance curves for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$

2.5.7 Implementation Issues

In the previous subsection, the performance improvement potential of the BiDFE (e.g., BAD, LC-BiDFE), over the conventional DFE and the NCDFE, was numerically demonstrated. In this subsection, the implementation issues or challenges posed by the BiDFE are discussed. As the goal of advanced signal processing at the receiver in a digital communication system is to improve the performance at very little additional cost, we discuss the complexity of the BiDFE and compare it with other equalizer architectures. Since, a BiDFE relies on non-causal processing of the received signal through the operation of time-reversal, latency or the overall detection delay of the BiDFE is of prime concern. The other issues that are discussed in this subsection are complexity, robustness to channel estimation errors and employability of the BiDFE to non-packet based streaming applications.

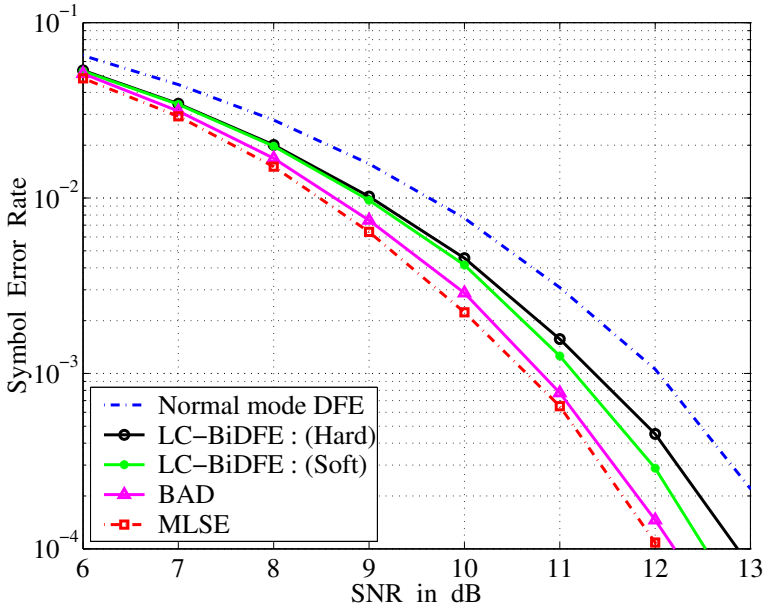

Figure 2.16: Performance of a Bidirectional Arbitrated DFE (BAD) for a symmetric channel with an impulse response $\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887]$

Latency

The time-reversal mode DFE can process the received signal only after the entire packet of data is received. So, the BiDFE will have an overall detection delay which is, at least, as large as the size of the packet. Such a detection delay may not be always acceptable, especially in certain applications like voice transmission, and is a drawback of the BiDFE structure. Hence, the choice of employing a BiDFE hinges critically on the latency requirements of the system. On the other hand, in some standards like GSM, the training sequence used for channel estimation is located at the center of the packet, namely as a mid-amble (see Figure 2.3). Hence, the detection delay for a GSM system with a conventional equalizer will be a little more than half the packet size. Furthermore, for GSM systems, the encoded data burst is interleaved and split in to four data packets before transmission. These four data packets have to be assembled at the receiver prior

to channel decoding. So, the delay of the BiDFE, may in fact, not be the limiting factor for such a system. In addition, when a hybrid ARQ protocol is used, a packet error results in a request for packet retransmission and the associated delay before the packet can be detected correctly. The performance gain offered by the BiDFE, on the other hand, will result in a decreased packet error rate (PER) and hence reduce the need for packet retransmission.

Streaming Applications

Although, the BiDFE structure has been designed for packet based transmission systems, it can be modified for streaming applications that lack a definite packetization structure. For a streaming system, the received signal can be broken down into overlapping virtual packets of some user-defined size. The virtual packets can then be processed using a BiDFE structure to improve performance at the cost of increased detection delay for the system. The reason for the packets to have some degree of overlap is to remove the edge effects that would be created due to the creation of these virtual packets.

Complexity

The BiDFE structure of Figure 2.9 uses two parallel DFE structures and, all things being equal, has a two-fold increase in complexity. For the performance comparisons provided in Figures 2.11 - 2.16, the BiDFE structure uses twice as many taps as either the normal mode DFE or the time-reversal mode DFE. In an attempt to make the comparison fair, Figure 2.17 illustrates[2] the MSE performance of the LC-BiDFE structure when compared with the normal mode DFE and the time-reversal mode DFE, both of which have the same total complexity as the LC-BiDFE. For instance, when a total number of

[2]See [19] for Matlab source code

14 taps are used, each of the constituent DFEs in the LC-BiDFE structure are allocated only 7 taps, while each of the normal mode and time-reversal mode DFEs, to which the LC-BiDFE performance is compared, are allocated 14 taps. In other words, the performance of a "double-length" DFE is compared with the performance of a LC-BiDFE structure that has two "single-length" constituent DFEs. The MSE plot is for the asymmetric channel with an impulse response $\mathbf{c}_1$ (see equation (2.66)) at an SNR of 15 dB. From Figure 2.17 it is evident that, for a reasonable choice on the total number of DFE taps, the performance improvement provided by the LC-BiDFE structure is similar to that shown in Figure 2.11. However, when the total number of taps in the LC-BiDFE is very small, the performance of each of the "single length" constituent DFEs in the BiDFE architecture is bad and this results in a poor overall performance.

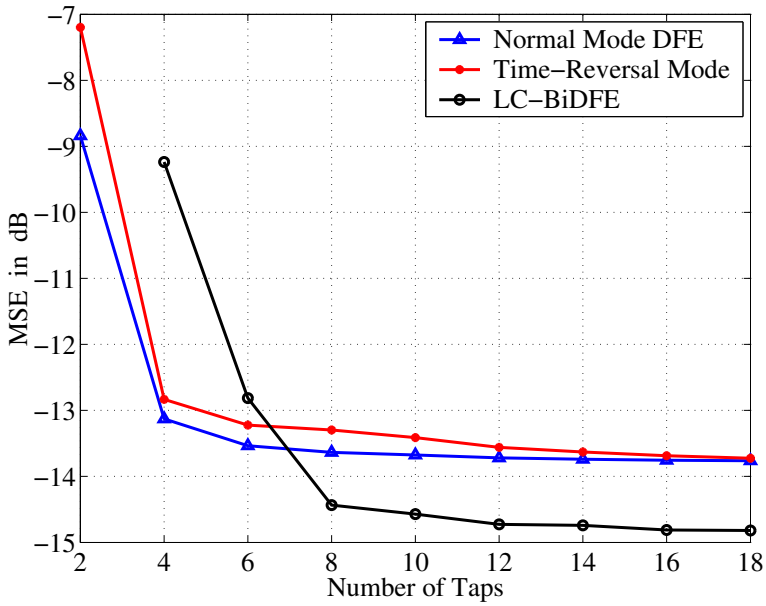

Figure 2.17: Comparative MSE performance with same total number of taps for a BiDFE and a DFE

The complexity of a linear equalizer, a conventional DFE, an LC-BiDFE, an NCDFE, BAD and MLSE are compared in Table 2.2. The number of multiplications and addi-

tions required for the computation of each symbol is tabulated. No attempt has been made to simplify the architecture of any of the schemes. For instance, when a BPSK source is equalized, the convolution operation of the FBF, which is typically a multiply-and-add operation can be replaced with a simple add/subtract operation[3]. However, these simplifications to the equalizer architecture are application specific and hence ignored here. For an NCDFE, L_p is the number of taps in the preliminary equalizer, for the BAD, $2K+1$ is the size of the window inside which the Euclidean distance metric is computed, and for the MLSE, $|\mathcal{A}|$ is the number of elements in the source alphabet.

Table 2.2: Complexity of equalizer structures

Equalizer Structure	**Number of Mult/Add Operations**
Linear Equalizer	L_f
DFE	$L_f + L_b$
LC-BiDFE	$L_{f1} + L_{b1} + L_{f2} + L_{b2}$
NCDFE	$L_p + L_f + L_b$
BAD	$L_{f1} + L_{b1} + L_{f2} + L_{b2} + 2L_c + 2(2K+1)$
MLSE	$(L_c+1)\|\mathcal{A}\|^{L_c}$

Channel Estimation

Although, we assume that the channel impulse response is perfectly known at the receiver (assumption 2.4), in practice the channel is estimated with the aid of the training sequence (recall Figure 2.3) and at times also the data. The estimated channel is typically imperfect and the severity of the estimation error is a function of the noise variance

[3]This would not be possible if a soft DFE [14] is employed

and the length of the training sequence. The channel estimation error can be incorporated in to the equivalent noise and it results in performance degradation. However, the degradation can be expected to have a comparable effect on performance of the normal mode DFE, the time-reversal mode DFE and the LC-BiDFE. As the BAD relies on channel estimates for the reconstruction based arbitration stage, it may experience an additional performance degradation.

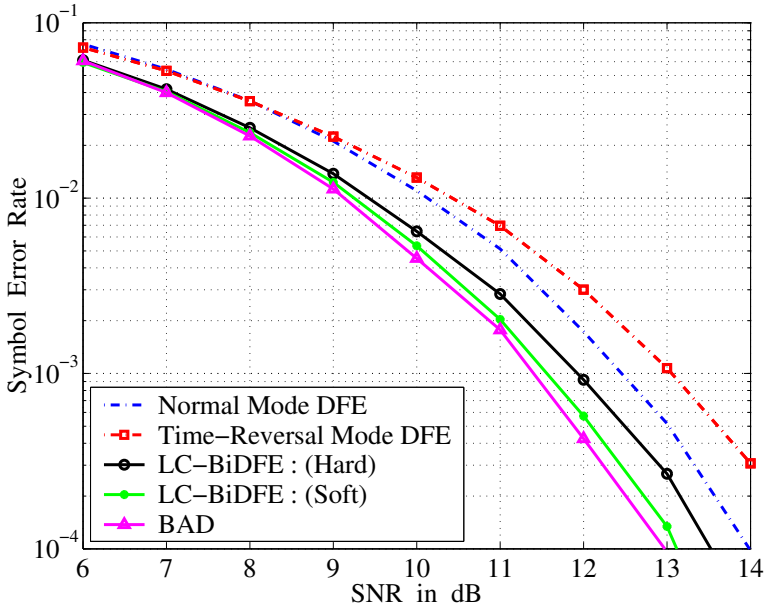

Figure 2.18: Effect of channel estimation on SER performance for an asymmetric channel with an impulse response of $\mathbf{c}_1 = [0.183 \;\; 0.916 \;\; 0.289 \;\; -0.183 \;\; 0.092 \;\; -0.046 \;\; 0.018]$

The robustness of the BiDFE was tested with a simulation example. The asymmetric channel impulse response $\mathbf{c}_1$, of equation (2.66), was considered. Assumption 2.4 was relaxed and the channel impulse response was estimated at the receiver using a Least-squares channel estimator [48]. A GSM training sequence [33], with a training length of 26 symbols, was used to aid the channel estimation. The SER performance curves[4]

[4]See [19] for Matlab source code

for a normal mode DFE, a time-reversal mode DFE, an LC-BiDFE (with and without soft decisions) and BAD are plotted in Figure 2.18. The normal mode and time-reversal mode DFEs were parameterized with 4 FFF taps and 3 FBF taps. The performance degradation exhibited by each equalizer structure in Figure 2.18, due to channel estimation errors at an SER of 10^{-3}, when compared to the SER curves of Figures 2.12 and 2.13 is tabulated in Table 2.3. The LC-BiDFE (with soft decisions) and BAD exhibit an additional performance degradation of about 0.25 dB - 0.3 dB at an SER of 10^{-3}, when compared with the normal mode DFE. However, even with channel estimation errors, the performance improvement of the LC-BiDFE (with soft decisions) is still about 0.9 dB.

Table 2.3: Performance degradation due to channel estimation

Equalizer Structure	**Performance Degradation**
DFE	0.8 dB
LC-BiDFE (Hard)	0.9 dB
LC-BiDFE (Soft)	1.06 dB
BAD	1.12 dB

2.6 Summary

A BiDFE structure that employs time-reversed processing of the received signal in conjunction with a normal mode DFE has been proposed in this Chapter. The proposed BiDFE structure has the twin advantage of simultaneously decreasing the gap from the MFB and mitigating error propagation (for instance BAD). The performance gain

provided by the LC-BiDFE structure has been demonstrated, both analytically and numerically. The performance of the BiDFE structures, LC-BIDFE and BAD, has been numerically compared with DFE enhancements of similar complexity, like the NCDFE. The performance gap from the MLSE for the various BiDFE structures is tabulated in Table 2.4 for the asymmetric channel impulse response $\mathbf{c_1}$ and an SER value of 10^{-3}. Implementation issues, namely latency, complexity and robustness to channel estimation errors, associated with the practical realization of the BiDFE structure have been discussed in detail.

Table 2.4: Performance comparison of equalizer etructures for $\mathbf{c_1}$

Equalizer Structure	**Peformance Gap from MLSE**
DFE	1.66 dB
LC-BiDFE (Hard)	1.07 dB
LC-BiDFE (Soft)	0.49 dB
BAD	0.27 dB

In the next Chapter, the performance limits of the LC-BiDFE will be investigated by considering an infinite length design and the MSE performance of the infinite length LC-BiDFE will be evaluated. Given a packet based system assumption and the employment of time-reversal, such an infinite length LC-BiDFE design seems counter-intuitive with its infinite latency. However, our interest is in characterizing the performance limits of the LC-BiDFE, by relaxing the finite length constraint. Hence, by letting the packet size tend to infinity asymptotically, we can obtain such an intuition.

Chapter 3

Infinite Length BiDFE

In this Chapter, an infinite length LC-BiDFE will be considered and the bound on the performance gain of this structure over the conventional DFE structure will be investigated. The performance of a finite length LC-BiDFE will asymptotically converge to that of the infinite length LC-BiDFE, when the total number of taps used by each of the constituent DFEs is increased. For a packet based communication system the idea of using an infinite length LC-BiDFE may seem counter-intuitive as the BiDFE architecture employs time-reversal of the received signal. However, our interest is in characterizing the performance limits of the LC-BiDFE by relaxing the finite length constraint, and thereby, obtain useful performance bounds. To employ an infinite length BiDFE the packet size also needs to be infinitely long; so we assume that $N \to \infty$.

This chapter is organized as follows. Section 3.1 describes the notation used in this chapter. Section 3.2 quantifies the advantage of the LC-BiDFE, over the conventional DFE, by computing the decrease in the gap from the MFB at high SNR scenarios. This motivates the tap optimization problem for the infinite length LC-BiDFE. In Section 3.3, the infinite length LC-BiDFE coefficients are optimized to minimize the overall MSE, rather than the MSE at the output of each the individual DFE streams, and Section 3.4

summarizes the results in this chapter.

3.1 System Model

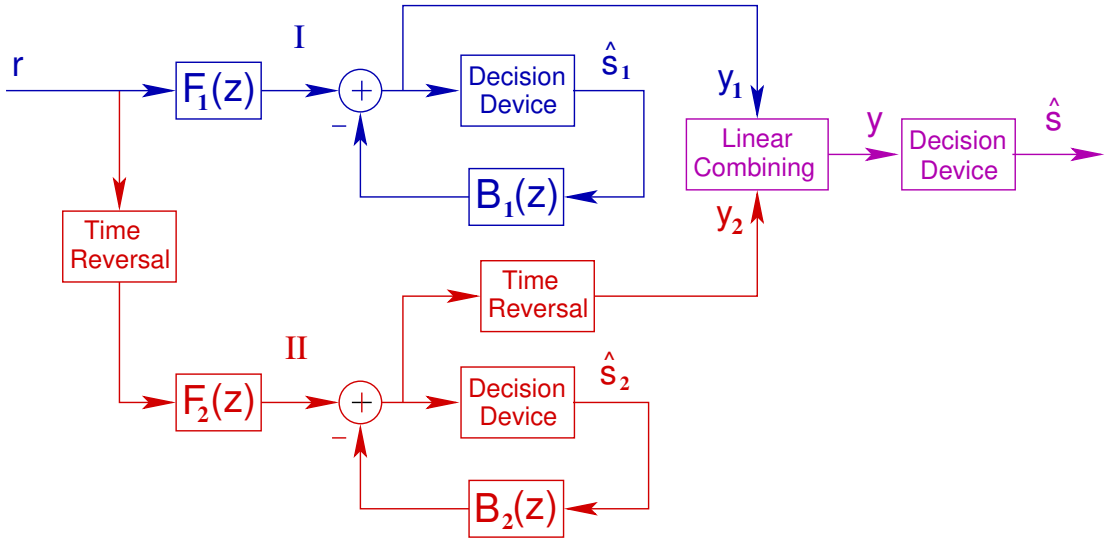

Figure 3.1: Structure of an LC-BiDFE.

The structure of an LC-BiDFE is depicted in Figure 3.1. Now, $F_1(z)$ and $F_2(z)$ represent the Z-transform of the FFF taps of the normal mode and time-reversal mode DFEs, respectively, and $B_1(z)$ and $B_2(z)$ are the Z-transform responses of the corresponding FBF taps. As stated in equation (2.52), the soft outputs of the normal mode and the time-reversal mode DFEs are combined using a memoryless weighted linear combiner, i.e.,

$$y(n) = \alpha y_1(n) + (1 - \alpha) y_2(n). \tag{3.1}$$

Also recall the MSE optimized combiner coefficient α of equation (2.54). In this chapter we invoke assumptions 2.1-2.6 and 2.8 of Chapter 2.5. For simplicity of notation, we also make the following additional assumption.

Assumption 3.1 *The channel impulse response has a unit norm, i.e.,* $||\mathbf{c}||^2 = 1$*, and does not have a spectral null, i.e., no roots on the unit circle.*

The assumption on the energy of the channel coefficients can be made without loss of generality, as it would not affect the output SNR of the equalizer, but only the absolute energy. The presence of spectral nulls is undesirable for the design of baud-spaced equalizers. Hence, we make the assumption on the absence of roots on the unit circle and this precludes only a small class of channels from the following analysis.

3.2 Performance of an Infinite Length LC-BiDFE

In this section, the gap from the MFB for the LC-BiDFE will be derived as a function of the channel impulse response for a high SNR scenario. Recall the definition of MFB from equation (2.5), namely

$$\gamma_{\text{mfb}} = \frac{||\mathbf{c}||^2}{\sigma_w^2}. \tag{3.2}$$

The infinite length MMSE-DFE tap coefficients derived in [70] were reviewed in subsection 2.2.2. The normal mode and time-reversal mode DFE taps are assumed to be parameterized with the MMSE-DFE coefficients. Recall the factorization of the channel impulse response from equation (2.33),

$$C(z) = c(0) \prod_{i=1}^{L_c-1} (1 - \beta_i z^{-1}) \ , \tag{3.3}$$

where β_i are the root locations of the channel impulse response polynomial. Then, the Z-transform of the time-reversed channel impulse response $\tilde{C}(z)$ is,

$$\begin{aligned}\tilde{C}(z) &\triangleq C(z^{-1}) \ , \\ &= c(0) \prod_{i=1}^{L_c} (1 - \beta_i z) \ , \\ &= c(L_c)(-z)^{L_c} \prod_{i=1}^{L_c} \left(1 - \frac{z^{-1}}{\beta_i}\right) \ . \end{aligned} \tag{3.4}$$

In equation (3.4), the second equality follows from the identity on the product of the roots of a polynomial,

$$\frac{c(L_c)}{c(0)} = (-1)^{L_c} \prod_{i=1}^{L_c} \beta_i \ . \tag{3.5}$$

Under the high SNR assumption, i.e., $\lambda \to 0$, the feedforward filter (FFF) for the normal mode and the time-reversal operations are given by (recall equations (2.23) and (2.25))

$$F_1(z) \approx \frac{1}{c(0)} \prod_{i=1, |\beta_i|>1}^{L_c} \frac{(\beta_i^* - z^{-1})}{\beta_i^*(1 - \beta_i z^{-1})} \ , \tag{3.6}$$

and

$$F_2(z) \approx \frac{(-z)^{-L_c}}{c(L_c)} \prod_{i=1, |\beta_i|<1}^{L_c} \frac{\beta_i(1 - \beta_i^* z^{-1})}{(\beta_i - z^{-1})} \ . \tag{3.7}$$

The feedback filters $B_1(z)$ and $B_2(z)$ are chosen such that the post-cursor ISI is perfectly canceled. Mathematically,

$$B_1(z) = \left\{ \prod_{i=1, |\beta_i|<1}^{L_c} (1 - \beta_i z^{-1}) \right\} \left\{ \prod_{i=1, |\beta_i|>1}^{L_c} (1 - (\beta_i^* z)^{-1}) \right\} - 1 \ , \tag{3.8}$$

and

$$B_2(z) = \left\{ \prod_{i=1, |\beta_i|<1}^{L_c} (1 - \beta_i^* z^{-1}) \right\} \left\{ \prod_{i=1, |\beta_i|>1}^{L_c} (1 - (\beta_i z)^{-1}) \right\} - 1 \ . \tag{3.9}$$

In [6], Ariyavisitakul derived the MSE and the output SNR of an infinite length DFE, under the ideal feedback assumption. These results will be briefly discussed here. The

MSE of the infinite length DFE for a high SNR scenario is

$$\begin{aligned} \text{MSE}_{\text{DFE}} &= E[|y_1(n) - s(n)|^2] \\ &= E\left[\left|\sum_k f_1(k)w(n-k)\right|^2\right] \cdot \end{aligned} \tag{3.10}$$

Since, the energy of an impulse response in time domain is the same as the constant term in the Z-transform polynomial expansion of its auto-correlation function [72], the MSE can be evaluated in the frequency domain as

$$\text{MSE}_{\text{DFE}} = \sigma_w^2 \left[F_1(z)F_1^*(z^{-1})\right]_0 \tag{3.11}$$

where the notation $[\Phi(z)]_0$ represents the constant term in the polynomial expansion of $\Phi(z)$, namely $\phi(0)$. With a few simple manipulations it can be shown that

$$\text{MSE}_{\text{DFE}} = \frac{\sigma_w^2}{c(0)c^*(0)} \left(\prod_{i=1,|\beta_i|>1}^{L_c} \beta_i\beta_i^*\right)^{-1} . \tag{3.12}$$

Ariyavisitakul also demonstrated that the MSE of both the normal mode and time-reversal mode DFEs are the same for the infinite length scenario.

Lemma 3.1 *The MSE of the output of the normal mode DFE and the time-reversal mode DFE are the same, under the high SNR assumption.*

Proof: To prove the above statement, it is sufficient to show that the energy of the feedforward filter of both the normal mode and time-reversal mode DFEs are the same, i.e., $||\mathbf{f}_1||^2 = ||\mathbf{f}_2||^2$. By using arguments similar to those in equations (3.10) - (3.12), we can show that

$$\left[F_2(z)F_2^*(z^{-1})\right]_0 = \frac{1}{c(L_c)c^*(L_c)} \left(\prod_{i=1,|\beta_i|<1}^{L_c} \beta_i\beta_i^*\right) . \tag{3.13}$$

Applying the result of equation (3.5) to equation (3.13),

$$[F_2(z)F_2^*(z^{-1})]_0 = \frac{1}{c(0)c^*(0)} \left(\prod_{i=1,|\beta_i|<1}^{L_c} \beta_i \beta_i^* \right) \left(\prod_{i=1}^{L_c} \beta_i \beta_i^* \right)^{-1} \tag{3.14}$$

$$= [F_1(z)F_1^*(z^{-1})]_0 \ . \qquad \blacksquare$$

The MMSE-DFE suffers a performance loss when compared to the MFB and the output SNR of the MMSE-DFE is

$$\gamma_{\text{DFE}} = \frac{c(0)c^*(0)}{\sigma_w^2} \left(\prod_{i=1,|\beta_i|>1}^{L_c} \beta_i \beta_i^* \right)^{-1} \ . \tag{3.15}$$

As both the normal mode and time-reversal mode streams yield the same MSE, an equal gain combining scheme, i.e., $\alpha = \frac{1}{2}$, is optimal for the diversity combining block of the LC-BiDFE of Figure 3.1. As the residual ISI components are absent under the high SNR approximation for the feedforward filters in equations (3.6) and (3.7), the overall MSE of the LC-BiDFE can be written as

$$\begin{aligned} \text{MSE}_{\text{BiDFE}} &= E[|y(n) - s(n)|^2] \\ &= E\left[\left| \sum_k \frac{\{f_1(k) + f_2(-k)\}\, w(n-k)}{2} \right|^2 \right] \ . \end{aligned} \tag{3.16}$$

The MSE can be similarly evaluated in the frequency domain as

$$\text{MSE}_{\text{BiDFE}} = \sigma_w^2 \left[F_{eq}(z) F_{eq}^*(z^{-1}) \right]_0 \tag{3.17}$$

where

$$F_{eq}(z) = \frac{F_1(z) + F_2(z^{-1})}{2} \ . \tag{3.18}$$

The performance gain provided by the LC-BiDFE over the DFE is given by,

$$\frac{\gamma_{\text{BiDFE}}}{\gamma_{\text{DFE}}} = \frac{[F_1(z)F_1^*(z^{-1})]_0}{\left[F_{eq}(z)F_{eq}^*(z^{-1})\right]_0} \ . \tag{3.19}$$

Lemma 3.2 *The gain in output SNR of the infinite length LC-BiDFE over the infinite length DFE, under a high SNR assumption is*

$$\frac{\gamma_{BiDFE}}{\gamma_{DFE}} = \frac{2}{1+\Re\{g(0)\}}\ , \tag{3.20}$$

where $\Re\{g(0)\}$ denotes the real part of $g(0)$, and the Z-transform of the response $g(k)$ is

$$G(z) = \frac{c(0)z^{Lc}}{c(L_c)}\left\{\prod_{i=1,|\beta_i|<1}^{L_c}\frac{\beta_i(\beta_i^*-z^{-1})}{(1-\beta_i(-z)^{-1})}\right\}\left\{\prod_{i=1,|\beta_i|>1}^{L_c}\frac{\beta_i^*(1-\beta_i z^{-1})}{(\beta_i^*-z^{-1})}\right\}\ . \tag{3.21}$$

Proof: From equation (3.18), the denominator of equation (3.19) can be rewritten as

$$\begin{aligned}[F_{eq}(z)F_{eq}^*(z^{-1})]_0 &= [\{F_1(z)+F_2(z^{-1})\}\{F_1^*(z^{-1})+F_2^*(z)\}]_0 \\ &= [F_1(z)F_1^*(z^{-1})+F_2(z^{-1})F_2^*(z)+F_1(z)F_2^*(z)+F_1^*(z^{-1})F_2(z^{-1})]_0\ . \end{aligned}\tag{3.22}$$

From Lemma 3.1 and equation (3.12),

$$[F_1(z)F_1^*(z^{-1}]_0 = [F_2(z^{-1})F_2^*(z)]_0 = \frac{1}{c(0)c^*(0)}\left(\prod_{i=1,|\beta_i|>1}^{L_c}\beta_i\beta_i^*\right)^{-1}\ . \tag{3.23}$$

Now define,

$$\begin{aligned}G(z) &\triangleq \frac{F_1^*(z^{-1})F_2(z^{-1})}{[F_1(z)F_1^*(z^{-1}]_0} \\ &= \frac{c(0)(-z)^{Lc}}{c(L_c)}\left\{\prod_{i=1,|\beta_i|<1}^{L_c}\frac{\beta_i(\beta_i^*-z^{-1})}{(1-\beta_i z^{-1})}\right\}\left\{\prod_{i=1,|\beta_i|>1}^{L_c}\frac{\beta_i^*(1-\beta_i z^{-1})}{(\beta_i^*-z^{-1})}\right\}\ . \end{aligned}\tag{3.24}$$

Let $g(0)$ denote the constant term in the discrete time domain expansion of $G(z)$. Then,

$$[F_1^*(z^{-1})F_2(z^{-1})]_0 = g(0)[F_1(z)F_1^*(z^{-1})]_0. \tag{3.25}$$

Since time-reversal of a discrete time sequence does not affect the entry with the index zero,

$$\begin{aligned}[F_1(z)F_2^*(z)]_0 &= [F_1^*(z^{-1})F_2(z^{-1})]_0^*, \\ &= g^*(0)[F_1(z)F_1^*(z^{-1})]_0. \end{aligned}\tag{3.26}$$

Hence, we can conclude that

$$[F_{eq}(z)F_{eq}^{*}(z^{-1})]_0 = [F_1(z)F_1^{*}(z^{-1})]_0[2 + 2\Re\{g(0)\}]. \tag{3.27}$$

■

3.2.1 Numerical Results

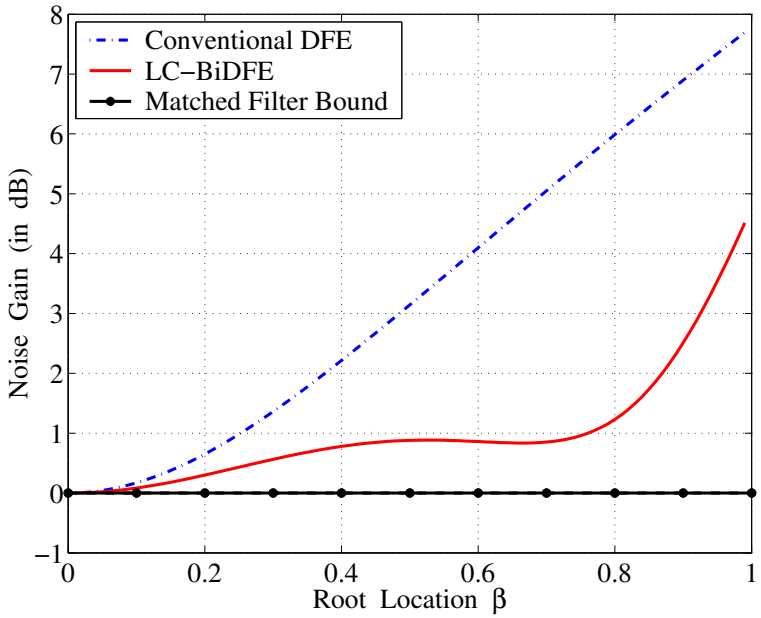

Figure 3.2: Performance gap from the matched filter bound for a symmetric 3-tap channel with root locations at $\left(\beta,\ \frac{1}{\beta}\right)$.

To illustrate the performance improvement that can be obtained by the use of an LC- BiDFE, we consider two test cases. Figure 3.2 illustrates[1] the gap from the MFB, for the conventional DFE and the LC-BiDFE, for a 3-tap real symmetric channel with root locations at β and $\frac{1}{\beta}$. These results correspond to the high SNR approximation that has been used in this section. It is clear that as the channel becomes severe, i.e., as

[1]See [19] for Matlab source code

the root locations move closer to the unit circle, the performance gain provided by the LC-BiDFE over the conventional DFE is more than 3 dB.

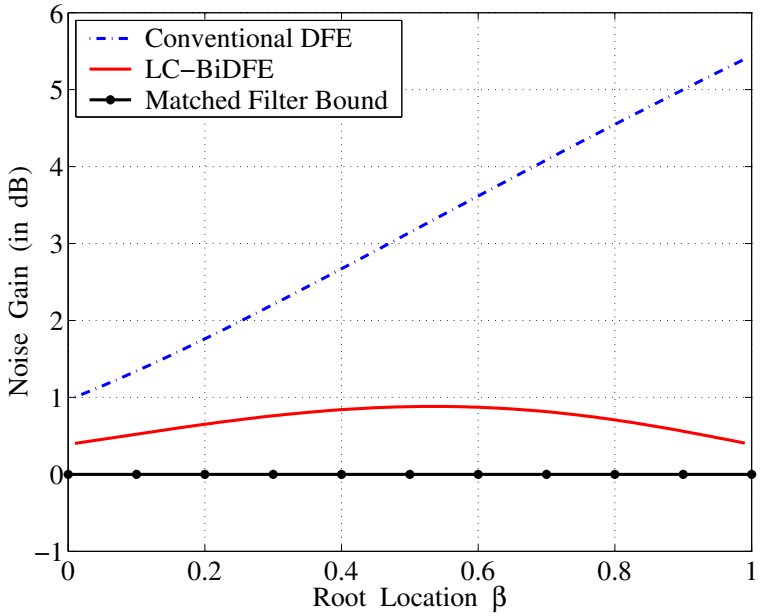

Figure 3.3: Performance gap from the matched filter bound for an asymmetric 3-tap channel with root locations at $\left(\beta,\ \frac{1}{2}\right)$.

Figure 3.3 illustrates[2] similar performance improvements for the LC-BiDFE, when applied to a 3-tap real asymmetric channel with root locations at β and $\frac{1}{2}$. Although the LC-BiDFE performs better than the MMSE-DFE, it still suffers from a small performance penalty when compared to the MFB. One possible reason is the fact that each of the constituent DFEs is chosen to minimize the MSE at its respective output. However, for the LC-BiDFE, the metric of interest is the overall MSE. Perhaps, if we attempt to optimize the taps of the LC-BiDFE to minimize the overall MSE, it might be possible to attain the MFB.

[2]See [19] for Matlab source code

3.3 LC-BiDFE Tap Optimization

In the analysis of Section 3.2 and in the examples corresponding to Figures 3.2 and 3.3, the DFE coefficients of each stream of the LC-BiDFE are optimized independently; an MMSE-DFE setting is chosen. In this section, we formulate the infinite length LC-BiDFE tap optimization problem to minimize the overall MSE of the LC-BiDFE. The resulting MSE would provide a bound on the potential performance improvements that can be achieved with this structure. It should be noted that the finite length results, which are more useful in practice, asymptotically converge to the infinite length results. A generalized formulation of the optimization problem involves determining the LC-BiDFE tap coefficients, namely $[\mathbf{f_1}, \mathbf{b_1}]$, $[\mathbf{f_2}, \mathbf{b_2}]$, and the coefficient of the diversity combiner α, that minimizes the overall MSE of the LC-BiDFE. Such an optimization problem is mathematically intractable. Hence, the LC-BiDFE structure of Figure 3.1 is simplified by introducing a front-end filter matched to the channel impulse response, namely $C^*(z^{-1})$, as shown in Figure 3.4.

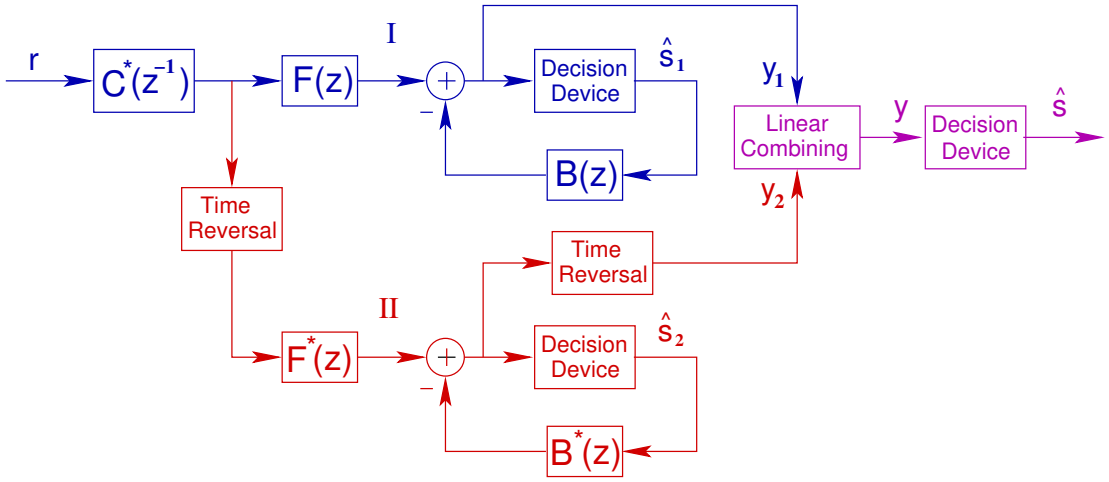

Figure 3.4: LC-BiDFE structure for tap optimization in an infinite length scenario

The advantage of the front-end matched filtering are two fold. Firstly, the combined

channel and matched filter impulse response $\Phi(z)$ is conjugate symmetric.

$$\Phi(z) = C(z)C^*(z^{-1}) \cdot \tag{3.28}$$

Secondly, as the additive noise sequence $w(n)$ is assumed to be white, the power spectral density of the noise at the output of the front-end matched filter is $\Phi(z)$. Let $[F(z), B(z)]$ denote the DFE tap coefficients of the normal mode DFE. As the time-reversal of $\Phi(z)$, due to its conjugate symmetry property, is $\Phi^*(z)$, a suitable choice for the tap coefficients of the time-reversal mode DFE is $[F^*(z), B^*(z)]$. It is straightforward to show that for this modified structure and the choice of DFE filter coefficients, both the normal mode and time-reversal mode streams will have the same value of MSE. Hence, an equal-gain combining scheme, with $\alpha = 1/2$, is optimal. The MSE minimization problem can now be cast in the form

$$(F^{\text{opt}}, B^{\text{opt}}) = \arg\min_{F,B} \mathrm{E}\left[|y(n) - s(n)|^2\right] \cdot \tag{3.29}$$

When optimizing the LC-BiDFE coefficients, a meaningful constraint to impose is for the FBF of each constituent DFE to remove all the post-cursor ISI in that stream. Otherwise, it is possible to indefinitely increase the MSE in each stream without affecting the overall MSE of the LC-BiDFE. Since, in reality the ideal feedback assumption has to be relaxed to allow for decision errors, this is critical. The following lemma will be used to optimize the taps of the infinite length LC-BiDFE.

Lemma 3.3 If any pair of polynomials $\{V(z), C(z)\}$ are such that $[V(z)]_0 = 1$ and $[C(z)C^*(z^{-1})]_0 = 1$, then

$$\left[\frac{V(z)V^*(z^{-1})}{C(z)C^*(z^{-1})}\right]_0 \geq 1 \tag{3.30}$$

and equality is attained if and only if $V(z) = C(z)C^*(z^{-1})$.

Proof: The term $[V(z)]_0$ can be evaluated using the well-known identity based on a convolution integral around the unit circle, i.e.,

$$[V(z)]_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(e^{j\theta})d\theta \cdot \tag{3.31}$$

Applying Cauchy-Schwartz inequality to the two continuous functions, $V(e^{j\theta})/C(e^{j\theta})$ and $C^*(e^{j\theta})$, we have

$$\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{V(e^{j\theta})}{C(e^{j\theta})}C(e^{j\theta})d\theta\right]^2 \leq \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{V(e^{j\theta})V^*(e^{j\theta})}{C(e^{j\theta})C^*(e^{j\theta})}d\theta\right] \times \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} C(e^{j\theta})C^*(e^{j\theta})d\theta\right] , \tag{3.32}$$

where equality is attained if and only if

$$\frac{V(e^{j\theta})}{C(e^{j\theta})} \propto C^*(e^{j\theta}) \tag{3.33}$$

or equivalently,

$$\frac{V(z)}{C(z)} \propto C^*(z^{-1}). \tag{3.34}$$

Since $[V(z)]_0 = 1$ and $[C(z)C^*(z^{-1})]_0 = 1$, equation (3.32) reduces to

$$\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{V(e^{j\theta})V^*(e^{j\theta})}{C(e^{j\theta})C^*(e^{j\theta})}d\theta\right] \geq 1 \tag{3.35}$$

with equality if and only if $V(z) = C(z)C^*(z^{-1})$. ■

Theorem 3.1 *The unbiased infinite length MMSE-BiDFE, under the ideal feedback assumption, attains the MFB.*

Proof: Consider the LC-BiDFE receiver structure illustrated in Figure 3.4. Under the constraint that the FBF perfectly removes the post-cursor ISI, we have

$$F(z)\Phi(z) = U(z) + 1 + B(z) \tag{3.36}$$

where $U(z)$ is a purely anti-causal response that represents the residual pre-cursor ISI, $B(z)$ is the purely causal FBF. The MSE of the LC-BiDFE is then given by

$$\begin{aligned}\text{MSE} = \frac{1}{4}\left[\{U(z)+U^*(z^{-1})\}^2\right]_0 \\ + \frac{\sigma_w^2}{4}\left[\Phi(z)\{F(z)+F^*(z^{-1})\}^2\right]_0 \cdot\end{aligned} \tag{3.37}$$

Since $U(z)$ is purely anti-causal, the first term can be simplified to

$$\frac{1}{4}\left[\{U(z)+U^*(z^{-1})\}^2\right]_0 = \frac{1}{2}[U(z)U^*(z^{-1})]_0 \tag{3.38}$$

and attains a minimum value of zero, if and only if $U(z)=0$. By using equation (3.36) and the conjugate symmetry property of $\Phi(z)$, the second term of equation (3.37) can be rewritten as

$$\begin{aligned}\frac{\sigma_w^2}{4}\left[\Phi(z)\{F(z)+F^*(z^{-1})\}^2\right]_0 = \\ \frac{\sigma_w^2}{4}\left[\frac{\{2+B(z)+B^*(z^{-1})+U(z)+U^*(z^{-1})\}^2}{\Phi(z)}\right]_0 \cdot\end{aligned} \tag{3.39}$$

Let us define

$$V(z) \triangleq 1 + \frac{B(z)+B^*(z^{-1})+U(z)+U^*(z^{-1})}{2} \cdot \tag{3.40}$$

Then, equation (3.39) simplifies to

$$\frac{\sigma_w^2}{4}\left[\Phi(z)\{F(z)+F^*(z^{-1})\}^2\right]_0 = \sigma_w^2\left[\frac{\{V(z)\}^2}{\Phi(z)}\right]_0 \cdot \tag{3.41}$$

From Lemma 3.3, the second term in equation (3.37) is minimized when $V(z) = \Phi(z)$. Hence, the MSE of equation (3.37) attains a minimum value when $U(z)=0$ and

$$1 + \frac{B(z)+B^*(z^{-1})}{2} = \Phi(z) \cdot \tag{3.42}$$

As the feedback filter $B(z)$ is purely causal, the optimal tap coefficients are given by,

$$b^{\text{opt}}(k) = 2\phi(k)\ , \qquad \forall\ k = 1, 2, \cdots, L_c \tag{3.43}$$

and the optimum FFF is

$$F^{\text{opt}}(z) = \frac{1 + B^{\text{opt}}(z)}{\Phi(z)} \ . \tag{3.44}$$

From equations (3.30), (3.42) and (3.44), the minimum MSE attained by this choice of BiDFE coefficients is

$$\text{MSE}_{\text{MMSE-BiDFE}} = \sigma_w^2 \tag{3.45}$$

and hence the maximum output SNR is

$$\gamma_{\text{MMSE-BiDFE}} = \frac{1}{\sigma_w^2} \ , \tag{3.46}$$

which is same as the MFB. In other words, the MSE optimized infinite length LC-BiDFE attains the MFB. ■

3.3.1 Relation to MSE optimized NCDFE

In the discussion on NCDFE in Chapter 2, subsection 2.3.1, it was stated that the optimal feedforward NCDFE taps are proportional to the matched filter response of the channel. Under the unit norm assumption on the channel, the constant of proportionality turns out to be unity. The feedback filter of the NCDFE is then given by,

$$b^{\text{ncdfe}}(k) = \phi(k) \ , \qquad \forall \ k = -L_c, \ldots, -1, 1, \ldots, L_c \tag{3.47}$$

Comparing equations (3.43) and (3.47), we see that although the length of the feedback filter in the MMSE-BiDFE is only half as long as the FBF of the NCDFE, the tap coefficients are twice in magnitude when compared to those of the NCDFE FBF taps.

3.3.2 Uniqueness of MMSE-BiDFE

In the earlier part of the section, the MMSE-BiDFE tap coefficients were optimized by restricting the structure of the receiver architecture. The resulting MMSE-BiDFE

was shown to attain the MFB, but are the LC-BiDFE filter settings that attain the MFB unique? In other words, are there other filter settings that would let the BiDFE attain the MFB. In this subsection, we show that the MSE optimization of the LC-BiDFE is not unique by considering a simple 2 tap example.

Example 3.1 *Consider a real 2-tap channel impulse response* $C(z) = a + \sqrt{1-a^2}z^{-1}$, *where* $|a| < 1$. *Then the MMSE-BiDFE filter settings that attain the MFB can be written from equations (3.43) and (3.44) as*

$$B^{opt}(z) = 2a\sqrt{1-a^2}z^{-1} \tag{3.48}$$

and

$$F^{opt}(z) = \frac{1 + 2a\sqrt{1-a^2}z^{-1}}{a\sqrt{1-a^2}z^{-1} + 1 + a\sqrt{1-a^2}z^{-1}} \cdot \tag{3.49}$$

On the other hand, consider the following filter settings.

$$F_1(z) = \frac{1}{a},\ B_1(z) = \frac{\sqrt{1-a^2}}{a}z^{-1} \tag{3.50}$$

and

$$F_2(z) = \frac{z^{-1}}{\sqrt{1-a^2}},\ B_2(z) = \frac{a}{\sqrt{1-a^2}}z^{-1}. \tag{3.51}$$

For these filter settings, the effective noise terms at the output of the normal mode and time-reversal mode DFEs are,

$$\eta_1(n) = \frac{w(n)}{a}, \quad \eta_2(n) = \frac{w(n+1)}{\sqrt{1-a^2}} \tag{3.52}$$

and

$$\sigma_1^2 = \frac{\sigma_w^2}{a^2}, \quad \sigma_1^2 = \frac{\sigma_w^2}{1-a^2}, \quad \rho = 0. \tag{3.53}$$

Hence, from equations (2.54) and (2.57), the optimal weighting factor α *and the MSE are given by*

$$\alpha = a^2, \quad \textit{MSE} = \sigma_w^2. \tag{3.54}$$

Hence, the filter settings of equation (3.50) and (3.51) also attain the MFB.

The 2-tap channel example 3.1 demonstrates that the settings of the infinite length LC-BiDFE, that attain the MFB, are not unique. In this chapter, the LC-BiDFE tap coefficients have been derived under the ideal feedback assumption. When the ideal feedback assumption is relaxed, the additional obstacle of error propagation has to be overcome. The performance loss due to error propagation increases with the magnitude of the feedback filter tap coefficients. If all the LC-BiDFE tap settings that attain the MFB are known, perhaps, a lower performance degradation can be achieved by choosing the tap setting with the least energy in its feedback filter taps.

3.4 Summary

In this chapter the performance of an infinite length LC-BiDFE with MMSE-DFE taps coefficients has been quantified. Although, the LC-BiDFE offers a significant performance gain from the conventional DFE structure, it still has a penalty from the MFB. In an attempt to decrease this gap further we formulated and solved the MSE optimization of the LC-BiDFE taps. The MMSE-BiDFE has been shown to attain the MFB, under the ideal feedback assumption. Since, in practice, only finite length filter realizations are possible, the optimization of the finite length LC-BiDFE taps will be considered in the next chapter.

Chapter 4

Finite Length BiDFE

In chapter 3, the optimization of the infinite length LC-BiDFE to minimize the MSE was considered. In this chapter, the tap optimization problem will be extended to an LC-BiDFE with a finite length constraint. In Section 4.1, a special class of channels, with a symmetric channel impulse response, is considered and the optimal unbiased DFE tap setting that minimizes the MSE at the output of the diversity combining block of the LC-BiDFE is derived. The effectiveness of the LC-BiDFE tap optimization is evaluated by numerical simulations. Although, the ideal feedback assumption is invoked for the tap optimization problem, the effect of decision feedback on the tap optimized LC-BiDFE is also studied via simulations. A drawback of the LC-BiDFE tap optimization, under decision feedback, is noted and two possible solutions are discussed. First, a modification of the MSE cost function to include an additional term, proportional to the energy in the FBF, is proposed in Section 4.2. Secondly, an iterative LC-BiDFE approach is proposed in Section 4.3. The LC-BiDFE tap optimization is extended to a general class of channels with an asymmetric channel impulse response in Section 4.4, and a summary of results is provided in Section 4.5.

4.1 MMSE-BiDFE for a Symmetric Channel

In Chapter 2, each of the normal mode and time-reversal mode DFEs were parameterized with the MMSE-DFE tap coefficients. However, these tap setting minimize the MSE at the output of each of the individual streams rather than the overall MSE. In this section, the filter taps of the finite length LC-BiDFE will be optimized to minimize the overall MSE of the system, namely $\mathrm{E}[\eta^2]$. In this section, we only consider channel impulse responses that are symmetric. The tap optimization for the case of an asymmetric channel will be addressed in Section 4.4.

For a symmetric channel $\mathbf{c}$, the block time-reversed impulse response, namely $\tilde{\mathbf{c}}$ is the same as $\mathbf{c}$. So, the same set of filter coefficients can be used for both the normal mode and the time-reversal mode DFE streams. Further more, this results in the property that both the normal mode and the time-reversal mode DFE streams have the same MSE performance. This leads to an equal-gain combining scheme, i.e., $\alpha = 1/2$, for the diversity combining block of Figure 3.1. Let δ denote the detection delay of each of the DFE streams. For the convenience of the reader, the $L_h \times L_f$ channel convolution matrix of equation (2.10) is restated here.

$$C = \begin{bmatrix} c(0) & 0 & \cdots & 0 \\ \vdots & c(0) & \cdots & \vdots \\ c(L_c-1) & \vdots & \cdots & 0 \\ 0 & c(L_c-1) & \cdots & c(0) \\ \vdots & 0 & \cdots & \vdots \\ 0 & \vdots & \cdots & c(L_c-1) \end{bmatrix} \tag{4.1}$$

Now, the combined channel feedforward impulse response is given by $\mathbf{h} = C\mathbf{f}$. From equation (2.16), to minimize the MSE under the ideal feedback assumption, the optimal FBF taps of the DFE should exactly cancel the L_b taps of $\mathbf{h}$ that follow the cursor h_δ

(recall equation (2.16)). Recall the $L_h \times L_h$ matrix V defined in equation (2.20). Define the vector $\mathbf{g}$ as

$$\mathbf{g} = VC\mathbf{f} - \mathbf{e}_\delta \tag{4.2}$$

where $\mathbf{e}_\delta$ is a column vector of length L_h and,

$$\mathbf{e}_\delta = [0_{1\times\delta}\ 1\ 0_{1\times L_h-\delta-1}]^T \tag{4.3}$$

Recall the expression for the residual ISI term of the noise sequence from equation (2.64), namely

$$\text{Residual ISI term} = \sum_k \left\{ \frac{g(k) + g(2\delta - k)}{2} \right\}^2 . \tag{4.4}$$

The residual ISI component can be expressed in vector form as

$$\text{Residual ISI term} = \frac{\mathbf{g}^T(S_1 + S_2)^T(S_1 + S_2)\mathbf{g}}{4} \tag{4.5}$$

where S_1 and S_2 are $(2L_h - 1) \times L_h$ matrices that perform shifting and time-reversal operations on the vector $\mathbf{g}$, prior to averaging. The matrices S_1 and S_2 are constructed as,

$$S_1 = \begin{bmatrix} 0_{(L_h-1-\delta)\times L_h} \\ I_{L_h\times L_h} \\ 0_{\delta\times L_h} \end{bmatrix}; \qquad S_2 = \begin{bmatrix} 0_{\delta\times L_h} \\ \tilde{I}_{L_h\times L_h} \\ 0_{(L_h-1-\delta)\times L_h} \end{bmatrix} \tag{4.6}$$

where $0_{\delta\times L_h}$ refers to a $\delta \times L_h$ dimensional zero matrix, I refers to an identity matrix and $\tilde{I}$ is a time-reversal matrix, i.e., a matrix with unit anti-diagonal entries and zero for all other entries.

$$\tilde{I}_{L_h\times L_h}(m,n) = \begin{cases} 1 & \text{if} \quad m+n = L_h+1 \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

Recall the expression for the noise gain term from equation (2.62),

$$\text{Noise gain term} = \sum_k \left\{ \frac{f(k) + f(2\delta + 1 - L_c - k)}{2} \right\}^2 \sigma_w^2. \tag{4.8}$$

This can be expressed in vector form as,

$$\text{Noise gain term} = \frac{\mathbf{f}^T(P_1+P_2)^T(P_1+P_2)\mathbf{f}\sigma_w^2}{4} \tag{4.9}$$

where P_1 and P_2 are $(L_h+L_f-1)\times L_f$ matrices that perform shifting and time-reversal operations on the vector $\mathbf{f}$, prior to averaging. The matrices P_1 and P_2 are constructed as shown below. If $2\delta+1-L_h \geq 0$ then,

$$P_1 = \begin{bmatrix} I_{L_f\times L_f} \\ 0_{(L_h-1)\times L_f} \end{bmatrix}; \qquad P_2 = \begin{bmatrix} 0_{(2\delta+1-L_h)\times L_f} \\ \tilde{I}_{L_f\times L_f} \\ 0_{(2L_h-2\delta-2)\times L_f} \end{bmatrix} \tag{4.10}$$

and if $2\delta+1-L_h < 0$ then,

$$P_1 = \begin{bmatrix} 0_{(L_h-2\delta-1)\times L_f} \\ I_{L_f\times L_f} \\ 0_{2\delta\times L_f} \end{bmatrix}; \qquad P_2 = \begin{bmatrix} \tilde{I}_{L_f\times L_f} \\ 0_{(L_h-1)\times L_f} \end{bmatrix}. \tag{4.11}$$

The optimal feedforward filter setting that minimizes overall MSE is given by,

$$\mathbf{f}_{\text{opt}} = \arg\min_{\mathbf{f}} \frac{1}{4}\Big[(VC\mathbf{f}-\mathbf{e}_\delta)^T(S_1+S_2)^T(S_1+S_2)(VC\mathbf{f}-\mathbf{e}_\delta) + \mathbf{f}^T(P_1+P_2)^T(P_1+P_2)\mathbf{f}\sigma_w^2\Big]. \tag{4.12}$$

The solution to the optimization problem of equation (4.12) is,

$$\mathbf{f}_{\text{opt}} = [(VC)^T(S_1+S_2)^T(S_1+S_2)VC + \sigma_w^2(P_1+P_2)^T(P_1+P_2)]^{-1} \cdot (VC)^T(S_1+S_2)^T(S_1+S_2)\mathbf{e}_\delta \tag{4.13}$$

The above MMSE filter setting for the BiDFE will be biased and can be unbiased by scaling $\mathbf{f}_{\text{opt}}$ by the reciprocal of the cursor.

$$\mathbf{f}_{\text{opt}}^{\text{bidfe}} = \frac{\mathbf{f}_{\text{opt}}}{h(\delta)} \tag{4.14}$$

and the optimum MMSE-BiDFE feedback filter is,

$$\mathbf{b}_{\text{opt}}^{\text{bidfe}} = MC\mathbf{f}_{\text{opt}}^{\text{bidfe}} \tag{4.15}$$

where M is the $L_b \times L_h$ matrix defined in equation (2.15).

4.1.1 Numerical Results

To test the effectiveness of the MMSE-BiDFE, in reducing the gap from the MFB, a real symmetric channel impulse response with an impulse response of

$$\mathbf{c_2} = [0.2887 \;\; 0.9129 \;\; 0.2887] \tag{4.16}$$

was considered. An FFF with 4 taps and an FBF with 3 taps were considered for each of the DFE streams. Recall from Figure 2.14, that the LC-BiDFE with MMSE-DFE taps suffers a loss of 0.6 dB from the MFB. Figure 4.1 compares[1] the MSE performance of a conventional DFE, an LC-BiDFE with MMSE-DFE taps, and an MSE optimized LC-BiDFE with the MFB. In the figure, the legend "suboptimal" refers to the case when MMSE-DFE tap settings are used, while "optimal" refers to the case when MMSE-BiDFE tap settings are used. From Figure 4.1, it can be seen that the MMSE-BiDFE almost closes the gap from the MFB, under the ideal feedback assumption.

In the presence of decision errors, one wonders if the LC-BiDFE with the MMSE-BiDFE tap settings can be expected to offer similar performance improvement over the LC-BiDFE with the MMSE-DFE filter settings? The answer to the above question, unfortunately, is a "no". The SER performance curves for the same channel example, assuming a BPSK source constellation, are shown in Figure 4.2. It can be seen in Figure 4.2, that the MMSE-BiDFE, instead of improving the performance, results in a performance worse than that of the "suboptimal" BiDFE. For instance, at an SER value of

[1] See [19] for Matlab source code

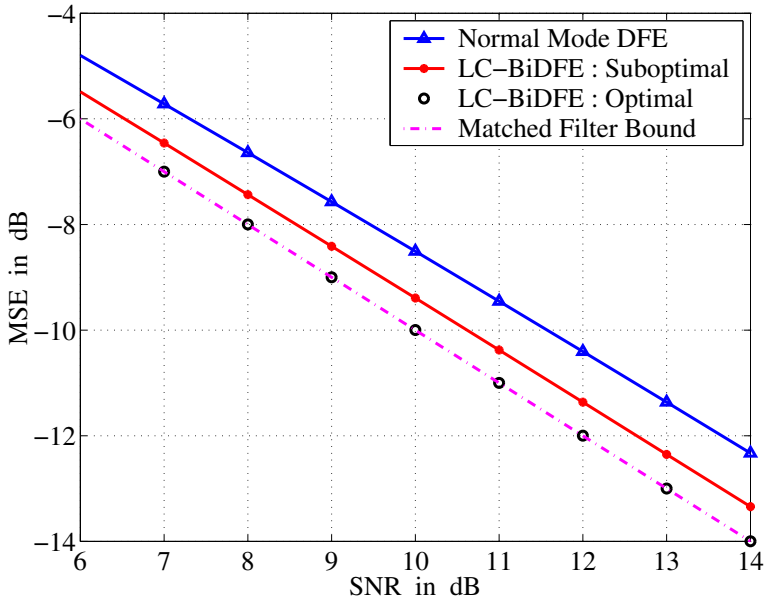

Figure 4.1: MSE performance of a finite length MMSE-BiDFE

10^{-3}, the MMSE-BiDFE (with soft feedback) has a performance loss of about 0.2 dB. In the presence of decision feedback, the SER curve is influenced not only by the decreased gap from MFB, but also by the error propagation effect. For an MMSE-BiDFE, the error propagation is larger because of the following two reasons:

- For the MMSE-BiDFE, although the overall MSE is lower, the MSE at the output of the normal mode and time-reversal mode DFE streams are higher, since neither is individually optimized, and this results in increased error propagation.

- The FBF tap settings are different for the MMSE-DFE and the MMSE-BiDFE. Simulations suggest that the FBF tap weights for the MMSE-BiDFE are usually larger in magnitude than the MMSE-DFE tap settings. As the error propagation gap increases with an increase in the magnitude of the FBF taps, this has an adverse effect on the SER performance of each of the constituent DFE streams, and hence the performance of the LC-BiDFE.

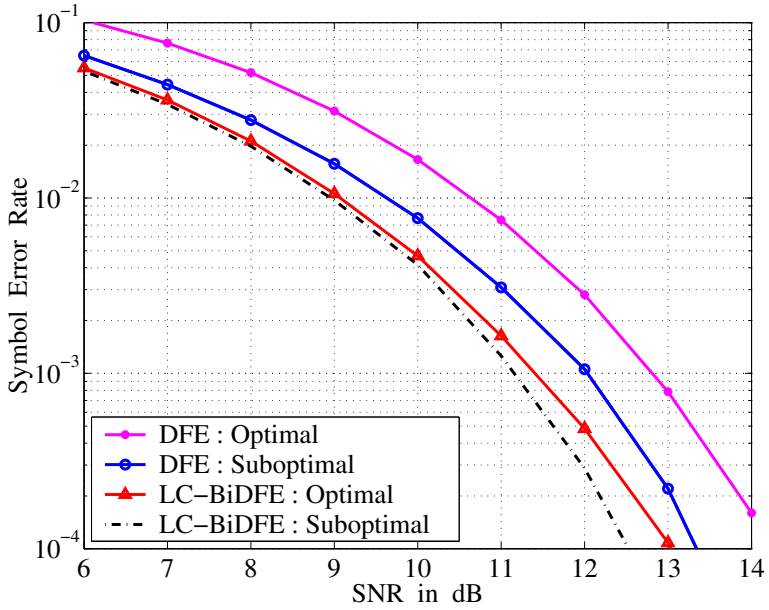

Figure 4.2: SER performance of a finite length MMSE-BiDFE

The effect of increased error propagation can be seen in Figure 4.2, where the SER performance of the normal mode DFE for the MMSE-BiDFE filter setting (labeled "Optimal" in the figure) is worse than the normal mode DFE performance for the MMSE-DFE filter setting (labeled "Suboptimal" in the figure) by about 0.8 dB for an SER value of 10^{-3}. In effect, although the MMSE-BiDFE tap settings offers a gain of about 0.6 dB in bridging the gap from the MFB, the error propagation gap increases by about 0.8 dB.

4.2 LC-BiDFE tap optimization with modified cost

In this section, we will modify the cost to be optimized in such a way that the optimal LC-BiDFE taps minimizes the effect of error propagation while minimizing the overall system MSE. There exist a few techniques in the literature that optimize the DFE taps to minimize the effect of error propagation. In [44, 45], Ghosh uses an approximate model

for error propagation to optimize the DFE taps by iteratively solving a set of coupled non-linear equations. However, this technique has a high computational complexity and cannot be readily applied to this problem. In [56], Kosut et al. modify the MSE cost function to include a norm constraint on the FBF taps and attempt to achieve a reduced error probability. Such an approach will be attempted here.

As the error propagation gap increases with an increase in the weights of the FBF tap coefficients, we include the norm of the FBF as the additional term in the MSE cost function of equation (4.12). The modified cost with the additional term is given by,

$$\mathbf{f}_{\text{opt}} = \arg\min_{\mathbf{f}} \frac{1}{4}\Big[(VC\mathbf{f} - \mathbf{e}_\delta)^T (S_1 + S_2)^T (S_1 + S_2)(VC\mathbf{f} - \mathbf{e}_\delta) + \mathbf{f}^T (P_1 + P_2)^T (P_1 + P_2)\mathbf{f}\sigma_w^2 + \xi (MC\mathbf{f}^T)(MC\mathbf{f})\Big] \quad (4.17)$$

where ξ is a user-defined weighting factor. When $\xi = 0$, we obtain the regular MMSE-BiDFE coefficients of equation (4.13), while other choices of ξ result in BiDFE taps with different MSE values and error propagation gap. As the weighting factor ξ is increased, settings with large value of FBF tap magnitudes incur a higher penalty. However, the overall system MSE corresponding to the optimal taps of the modified cost will be higher. The solution to the optimization of the modified cost function of equation (4.17) is,

$$\mathbf{f}_{\text{opt}} = \big[(VC)^T (S_1 + S_2)^T (S_1 + S_2) VC + \sigma_w^2 (P_1 + P_2)^T (P_1 + P_2) + \xi (MC)^T (MC) \big]^{-1} \cdot (VC)^T (S_1 + S_2)^T (S_1 + S_2)\mathbf{e}_\delta \quad (4.18)$$

As in section 4.1, the DFE taps can be unbiased as follows.

$$\mathbf{f}_{\text{mopt}}^{\text{bidfe}} = \frac{\mathbf{f}_{\text{opt}}}{h(\delta)} \quad (4.19)$$

and the "modified" MSE optimized LC-BiDFE feedback filter is,

$$\mathbf{b}_{\text{mopt}}^{\text{bidfe}} = MC\mathbf{f}_{\text{mopt}}^{\text{bidfe}} \quad (4.20)$$

where M is the $L_b \times L_h$ matrix defined in equation (2.15).

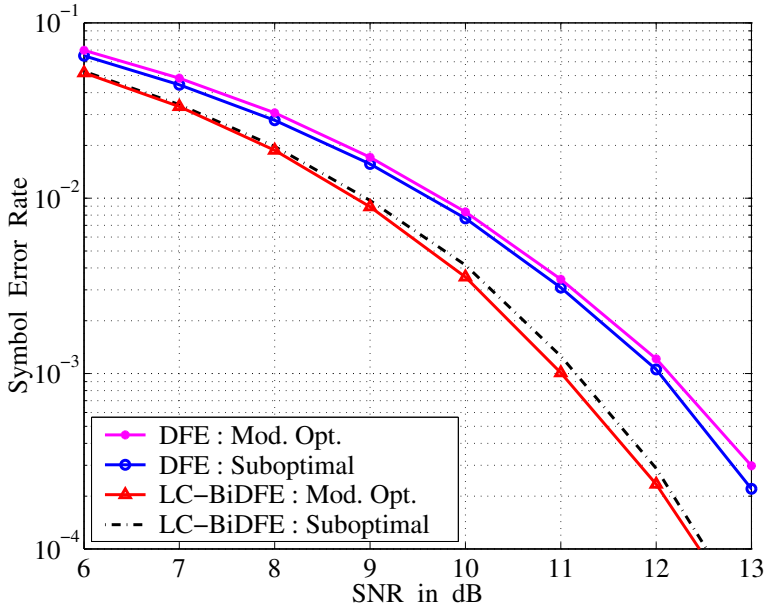

Figure 4.3: SER performance of a finite length "modified" MMSE-BiDFE

Table 4.1: ξ vs. SNR for the LC-BiDFE tap optimization with the modified cost function for $\mathbf{c}_2$

SNR	6	7	8	9	10	11	12	13	14
ξ	0.27	0.23	0.19	0.16	0.13	0.11	0.09	0.07	0.055

The effectiveness of including the norm term in the MSE cost function can be evaluated by considering the SER performance curves for the symmetric channel $\mathbf{c}_2$ of equation (4.15). Figure 4.3 compares[2] the performance of a BiDFE (with soft feedback) for the MMSE-DFE tap setting and the "modified" optimal tap setting of equations (4.19) and (4.20). The ξ values chosen for this simulation are tabulated in Table 4.1. Unlike, Figure 4.2, we observe a performance improvement, although marginal, for the "modified" optimal BiDFE tap settings. Also notice that the normal mode DFE performance is

[2]See [19] for Matlab source code

only marginally worse when compared to the MMSE-DFE (labeled "DFE: Suboptimal" in the figure) tap settings. The inclusion of the norm term ensures that the FBF tap coefficients are comparable in magnitude to those of the MMSE-DFE tap weights. This has been illustrated by plotting the FBF tap weights for the MMSE-DFE, the MMSE-BiDFE and the BiDFE optimized with the "modified" cost in Figure 4.4. This plot corresponds to an SNR value of 10 dB and $\xi = 0.2$. The FBF tap weights of the MMSE-BiDFE are about 60 % as large as the MMSE-DFE taps and hence result in an increased error propagation gap.

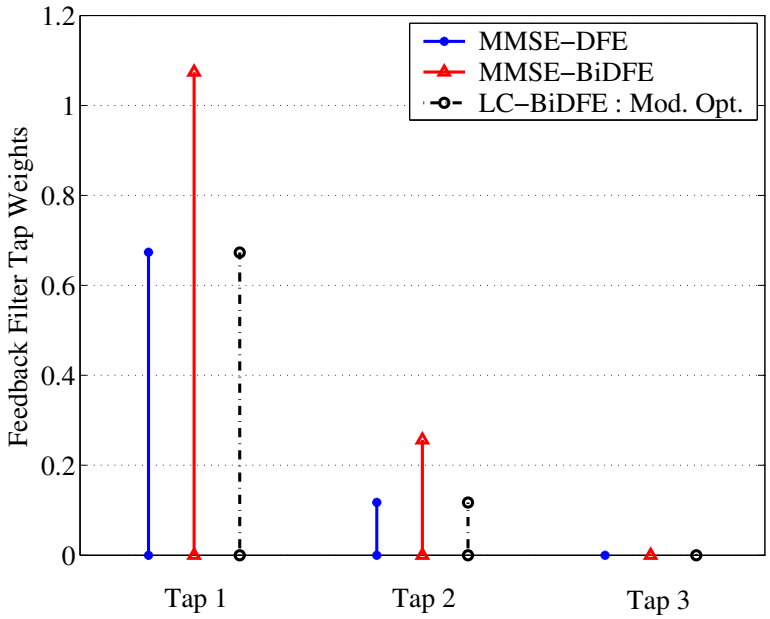

Figure 4.4: Comparison of feedback filter tap weights

However, one drawback of this technique is the choice of the user-defined parameter ξ. The ξ values used for the simulation example of Figure 4.3 are tabulated in Table 4.1. In the above example, the value of ξ was chosen based on simulations to maximize the performance gain of the BiDFE. Hence, the difficulty in choosing ξ renders this approach impractical.

4.3 LC-BiDFE with Iteration

The effect of increased error propagation when using the MMSE-BiDFE tap coefficients was illustrated in Figure 4.2. In this section an iterative solution is proposed to address this issue. The idea is to equalize the received signal using an LC-BiDFE with MMSE-DFE tap coefficients to provide an initial estimate of the transmitted symbols. In the second iteration, the received signal is re-equalized using an MMSE-BiDFE. During this iteration, the estimates of the first iteration are used to cancel the post-cursor ISI in both the normal mode and time-reversal mode DFEs and the outputs of the two modes are combined as before. The SER performance curve of the proposed iterative LC-BiDFE scheme is illustrated[3] in Figure 4.5 for the channel $\mathbf{c}_2$. The performance improvement is similar to that of Figure 4.3, and this technique does not have practical difficulties like choosing ξ encountered in Section 4.2. However, this performance improvement comes at the cost of a two-fold increase in complexity. An additional simulation example, demonstrating the effectiveness of this approach in offering a marginal improvement in SER performance, is provided in Appendix A.

4.4 MMSE-BiDFE for an Asymmetric Channel

The MSE optimization of the LC-BiDFE for an asymmetric channel is more difficult as it involves the joint optimization of both the DFE filter taps and the weighting factor α. Hence, we propose a finite length equivalent of the structure proposed in Figure 3.4 of section 3.3. In this structure, the received signal is processed with a front-end filter matched to the channel impulse response. Hence, the effective response seen by the normal mode and time-reversal mode DFEs will be symmetric. Furthermore, the auto-

[3] See [19] for Matlab source code

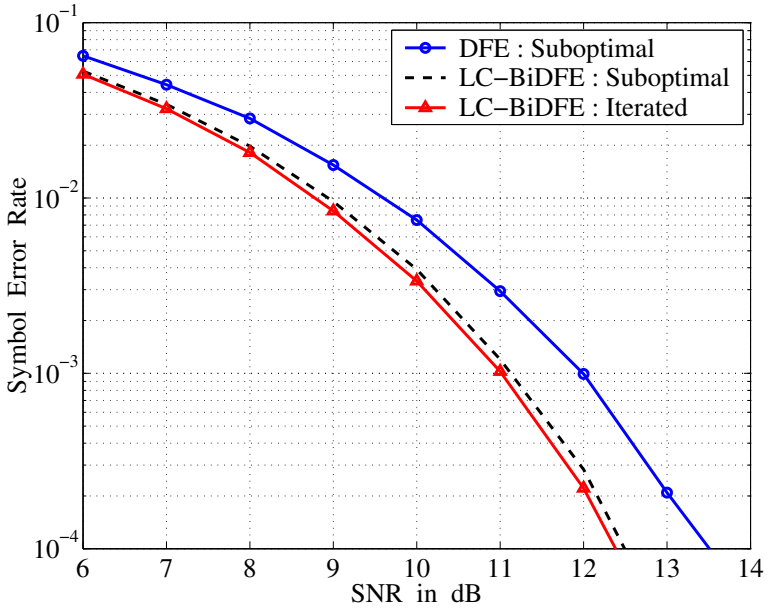

Figure 4.5: SER performance of an iterative finite length LC-BiDFE

correlation of the noise filtered through the channel impulse response is also symmetric. Hence, the same tap coefficients $(\mathbf{f}, \mathbf{b})$ can be used for both the normal mode and time-reversal mode DFE streams. An equal gain combining scheme, $\alpha = 1/2$, is optimal for the diversity combining block.

Let ϕ be the auto-correlation vector of the channel impulse response with a length of $L_\phi = 2L_c - 1$,

$$\phi(l) = \sum_{k=0}^{L_c-1} c(k)c(k + L_c - 1 - l), \quad \forall\ 0 \leq l < 2L_c - 1. \tag{4.21}$$

Let Φ represent the $L_h \times L_f$ channel convolution matrix, where $L_h = L_\phi + L_f - 1$. Let $\mathbf{h} = \Phi\mathbf{f}$ denote the combined channel, front-end matched filter and feedforward filter impulse response. Let us now define the $(L_c + L_f - 1) \times L_f$ channel convolution

matrices C (see equation (4.1)) and $\tilde{C}$, with

$$\tilde{C} = \begin{bmatrix} c(L_c-1) & 0 & \cdots & 0 \\ \vdots & c(L_c-1) & \cdots & \vdots \\ c(0) & \vdots & \ddots & 0 \\ 0 & c(0) & \ddots & c(L_c-1) \\ \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & \cdots & c(0) \end{bmatrix}. \tag{4.22}$$

The residual ISI term will have the same form as that of equation (4.5), except for the use of the convolution matrix Φ, instead of C. The noise term, on the other hand, will be a little different due to the front-end matched filter through which the additive white noise $w(n)$ is filtered. Incorporating the front-end matched filter, the feedforward filter optimization problem to minimize the MSE can be cast in the following form.

$$\mathbf{f}_{\text{opt}} = \arg\min_{\mathbf{f}} \frac{1}{4}\Big[(V\Phi\mathbf{f} - \mathbf{e}_\delta)^T (S_1+S_2)^T (S_1+S_2)(V\Phi\mathbf{f} - \mathbf{e}_\delta) + \mathbf{f}^T (P_1\tilde{C} + P_2 C)^T (P_1\tilde{C} + P_2 C)\mathbf{f}\sigma_w^2 \Big] \tag{4.23}$$

where the shifting and time-reversal matrices S_1, S_2 are the same as those defined in equation (4.6). The matrices P_1 and P_2 are $(L_h + L_c + L_f - 2) \times (L_c + L_f - 1)$ matrices that perform shifting and time-reversal operations on the vectors $\tilde{C}\mathbf{f}$ and $C\mathbf{f}$, respectively. If $2\delta + 1 - L_h \geq 0$ then,

$$P_1 = \begin{bmatrix} I_{(L_c+L_f-1)\times(L_c+L_f-1)} \\ 0_{(L_h-1)\times(L_c+L_f-1)} \end{bmatrix} \quad ; \quad P_2 = \begin{bmatrix} 0_{(2\delta+1-L_h)\times(L_c+L_f-1)} \\ \tilde{I}_{(L_c+L_f-1)\times(L_c+L_f-1)} \\ 0_{(2L_h-2\delta-2)\times(L_c+L_f-1)} \end{bmatrix} \tag{4.24}$$

and if $2\delta + 1 - L_h < 0$ then,

$$P_1 = \begin{bmatrix} 0_{(L_h-2\delta-1)\times(L_c+L_f-1)} \\ I_{(L_c+L_f-1)\times(L_c+L_f-1)} \\ 0_{2\delta\times(L_c+L_f-1)} \end{bmatrix} \quad ; \quad P_2 = \begin{bmatrix} \tilde{I}_{(L_c+L_f-1)\times(L_c+L_f-1)} \\ 0_{(L_h-1)\times(L_c+L_f-1)} \end{bmatrix} \tag{4.25}$$

The optimal FFF tap setting of equation (4.23) is

$$\mathbf{f}_{\text{opt}} = [(V\Phi)^T(S_1+S_2)^T(S_1+S_2)V\Phi + \sigma_w^2(P_1\tilde{C}+P_2C)^T(P_1\tilde{C}+P_2C)]^{-1} \cdot (V\Phi)^T(S_1+S_2)^T(S_1+S_2)\mathbf{e}_\delta. \quad (4.26)$$

As before, these taps can be unbiased by multiplying with the reciprocal of the cursor term. A drawback of the approach outlined here is the fact that a front-end matched filter is necessary in addition to the normal mode and time-reversal mode DFEs. Hence, this increases the total complexity of the receiver.

4.5 Summary

In this chapter, the optimal finite length MMSE-BiDFE taps were derived under the ideal feedback assumption for both symmetric and asymmetric channel impulse responses. Although, the MMSE-BiDFE taps decrease the gap from the MFB, when compared to the LC-BiDFE with MMSE-DFE tap settings, these gains do not translate to the SER performance curves in the presence of decision feedback. The use of a modified cost function, incorporating an addition term based on the norm of the FBF taps, was proposed to ensure that performance gains, even if marginal, can also be obtained in the presence of decision feedback. However, the difficulty in appropriately choosing the weighting factor, ξ, for the norm term makes this approach impractical. An alternate method incorporating an iterative LC-BiDFE structure has also been proposed, and the performance improvements are similar to those obtained with the modified cost function. In the next Chapter, we consider the extension of the BiDFE approach for a multi-input-multi-output (MIMO) channel equalization problem.

Chapter 5

BiDFE for MIMO Channel Equalization

In the earlier chapters, equalization of a single-input single-output (SISO) system was considered. In this chapter, the focus will be on multiple-input multiple-output (MIMO) systems. Channels with MIMO characteristics occur frequently in modern communication systems. A MIMO system is typically characterized by the transmission of multiple-input signals through a linear, dispersive, noisy channel and results in multiple-output signals at the receiver. The received signals are composed of a sum of several transmitted signals corrupted by ISI, co-channel interference (also known as multi-access interference, i.e., MAI) and noise. Examples of MIMO channels include TDMA digital cellular systems with multiple transmitter and receiver antennas, wide-band asynchronous CDMA systems, dually polarized radio channels and magnetic recording channels. Even in a single user communication system, there exist scenarios where a MIMO modeling of the system proves to be useful. One such example is a pulse amplitude modulated (PAM) cyclostationary sequence in the presence of ISI [30].

Recent antenna technology advances have made it possible to support multiple trans-

mit and receive antennas at the terminal [38, 90]. Particularly for large size data terminals such as laptops, it is possible to have up to four integrated antennas with sufficient spacing so that the correlation of the transmitted/received signals across the antennas is small. Phased array antennas and widely-spaced diversity antennas are two ways to use multiple antennas to provide improved spectral efficiency. In the first case a narrow beam directed towards the terminal is formed by transmitting the same signal, appropriately weighted in amplitude and phase, from each antenna element, while in the later case different signals are transmitted from the different antennas in order to take advantage of scattering through space-time coding. Space-Time Coding (STC) can be used in different ways: some use the additional antenna elements to provide diversity gain (e.g. [84]), while other techniques, such as BLAST (Bell Labs Layered Space-Time) [39], achieve higher data rates by transmitting independent data streams through each transmit antenna element.

The BLAST scheme was initially proposed in [37] to increase the spectral efficiency in a flat fading wireless environment. If the number of receive antennas is greater than or equal to the number of transmit antennas, it is possible to separate the signals from the different transmit antennas. Hence, one can potentially send independent information from the different transmit antennas and thereby increase the data rate of the link. The extension of such a scheme to a frequency- selective fading environment is illustrated in Figure 5.1.

5.1 System Model

A MIMO channel model with M inputs and P outputs is considered. An independent data stream is assumed to be transmitted through each input of the MIMO channel. The

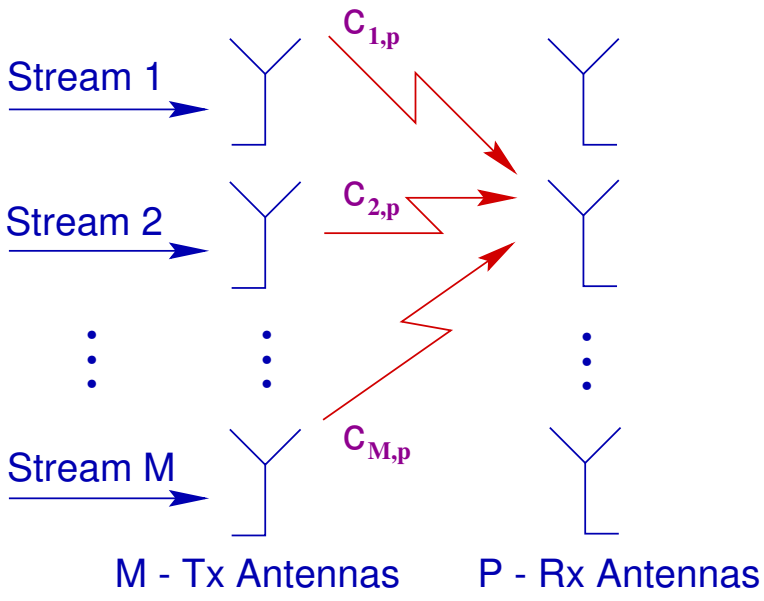

Figure 5.1: BLAST scheme for a multi-element antenna system

received sampled signal vector at the p-th output of the channel during the n-th symbol period can be expressed by the discrete-time model,

$$r_p(n) = \sum_{m=1}^{M} \sum_{k=0}^{L_c-1} c_{m,p}(k) s_m(n-k) + w_p(n) \tag{5.1}$$

where $\mathbf{c}_{m,p} \triangleq [c_{m,p}(0), \ \ldots, \ c_{m,p}(L_c-1)]^T$ is the impulse response of the channel between the m-th input and the p-th output, $s_m(n)$ is the transmitted symbol from the m-th input and $w_p(n)$ is the additive noise at the p-th output. Each of the channel impulse responses $\mathbf{c}_{m,p}$ is assumed to be time-invariant (assumption 2.1), FIR with L_c taps (assumption 2.2) and known at the receiver (assumption 2.4). The noise sequences $w_p(n)$ are assumed to be white, uncorrelated with each other and the source sequences, and of variance σ_w^2 (assumption 2.3).

5.1.1 Multichannel Matched Filter Bound

Unlike the SISO system, each transmitted information stream encounters a single-input multiple-output (SIMO) channel. Hence, the MFB for each stream has to be appropriately modified to reflect this. If the multiple-output channel for each stream is known, then a spatio-temporal matched filtering operation can be performed at the receiver to obtain a single-channel equivalent. Such a multichannel MFB was proposed by Slock and de Carvalho in [82]. For user i, the MFB is given by,

$$\gamma^i_{\text{mfb}} = \frac{\sigma_s^2 \sum_{p=1}^{P} \sum_{k=0}^{L_c-1} |c_{i,p}(k)|^2}{\sigma_w^2}. \tag{5.2}$$

As in the SISO case, the MFB can be attained in the MIMO scenario under the following conditions,

1. When only one symbol, pertaining to the stream i, is transmitted with no symbols transmitted for all other streams. This results in no ISI and no CCI and the matched filter receiver is the optimal detector for stream i.

2. If all transmitted symbols of all streams in the packet, except the symbol of interest of stream i, are known. In this case, the ISI and CCI components affecting the symbol of interest at the output of the matched filter receiver can be canceled perfectly.

5.2 MIMO Equalization

As in the case of a SISO system, the MIMO channel can be equalized using sequence estimation algorithms, such as the MLSE [34]. Although sequence estimators have superior performance when compared to symbol-by-symbol estimators, they have a high computational complexity. The MIMO MLSE scheme requires $|\mathcal{A}|^{ML_c}$ states for

the detection of a source selected from an alphabet set $\mathcal{A}$ transmitted using M antennas through a channel with a delay spread of L_c symbols. For instance, in the third-generation wireless TDMA proposal [40] 8-PSK modulation is used. A typical urban EDGE channel (including the transmit pulse shape) has a delay spread of at least 4 symbols. Even with two transmit antennas, the number of MLSE states required is $8^8 = 16777216$.

The low complexity symbol-by-symbol detector that is considered here is the MIMO DFE [30, 22, 2]. In [30], Duel Hallen derived the MMSE-DFE for the MIMO channel equalization problem, while in [2], the MIMO extension to the MMSE-DFE was derived under a finite length constraint. The MIMO DFE employs a feedforward filter $\mathbf{F}$ and a purely causal feedback filter $\mathbf{B}$. The block diagram of a MIMO DFE is illustrated in Figure 5.2. Typically, the MIMO feedforward filter suppresses the pre-cursor ISI and CCI (co-channel interference), while the MIMO feedback filter cancels the post-cursor ISI and the cursor/post-cursor CCI using the symbol decisions. As in the case of a SISO DFE, the MIMO DFE suffers from error propagation and a gap from the MFB. One aspect that makes the MIMO DFE different from the SISO DFE is the many variations that are possible in the realization of the MIMO DFE. These variations arise from the different ways in which the task of canceling the post-cursor CCI and the cursor CCI of the previously detected streams is partitioned between the FFF and the FBF.

Scenario 1

In this case, the FBF is used only to cancel the post-cursor ISI of the user (stream) of interest. The CCI from the interfering users are to be suppressed by the FFF. This structure is natural when only one of the streams is to be detected at the receiver, for example, detection at the mobile in downlink transmission for a wideband CDMA system.

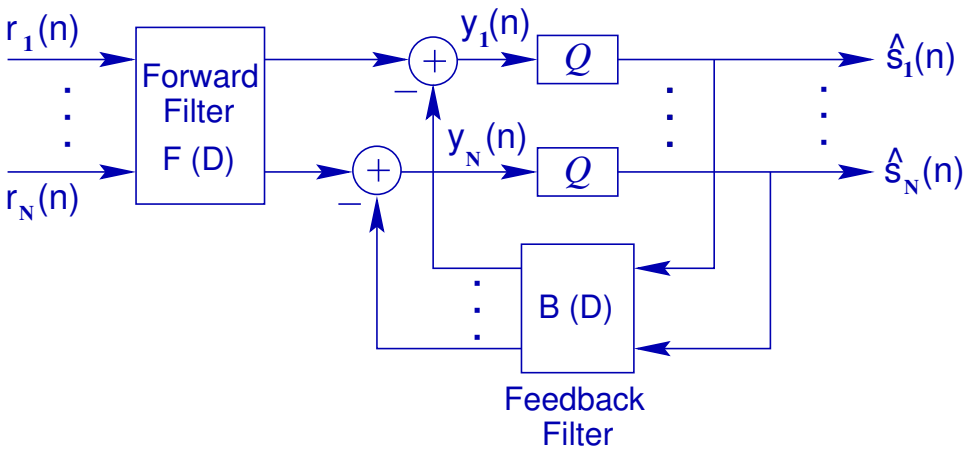

Figure 5.2: Structure of a MIMO DFE

Scenario 2

This assumes that all the users are equalized at the receiver and so the FBF has access to the post-cursor CCI from all interfering users in addition to the post-cursor ISI. Hence, the post-cursor CCI components from all detected streams are canceled using the FBF. The FFF is only used to suppress the pre-cursor ISI, pre-cursor and cursor CCI components.

Scenario 3

In scenario 2, if we assume that the users are detected in a certain order, then the FBF can be potentially modified to cancel not only the post-cursor CCI, but also the cursor CCI of the interfering users that have been detected prior to the user of interest. In such a scheme, while canceling the contribution of the cursor/post-cursor ISI of the detected users, the corresponding feedback filter section will have an additional tap when compared to the feedback filter sections of the users that have not been detected yet. The MMSE, under the ideal feedback assumption, improves progressively for each of the

structures described in scenarios 1-3 (see [2] for proof). Intuitively, when compared to scenario 1, as scenario 3 assumes that more of the interfering symbols (both ISI and CCI) are known, it has a lesser penalty from the MFB and a better performance. However, the severity of error propagation, i.e., when the decision-error-free assumption is violated, also progressively increases from scenario 1 to scenario 3.

5.3 BiDFE Extension to MIMO DFE

In this section the extension of the idea of a bidirectional DFE to MIMO channel equalization is considered. The MIMO DFE structure of scenario 3 is considered for this extension. The reason being that the MIMO DFE structure of scenario 3 affords the possibility of not only time-reversal but also user re-ordering. In other words, one could use two MIMO DFE structures, each with a different ordering for the detection of the users and combine the two MIMO DFE outputs to improve performance. The idea of employing user re-ordering to improve performance was proposed by Barriac and Madhow in [15] for a successive interference cancellation based multiuser detector for a CDMA system. The MIMO system model considered in [15] was assumed to be free from time dispersion, namely flat fading channels. Here, the use of time-reversal in conjunction with user re-ordering is proposed for channels with time dispersion.

Consider a MIMO BiDFE structure similar to the SISO BiDFE illustrated in Figure 2.9. The vector of received signals are processed using a normal mode MIMO DFE. Let $\mathbf{ord}_1 = [s_1(n),\ s_2(n),\ \ldots,\ s_M(n)]$ be the vector that denotes the order of detection for the normal mode MIMO DFE. The received signals are now block time-reversed and equalized using a time-reversal mode MIMO DFE. Further, the order of detection for time-reversal mode MIMO DFE, $\mathbf{ord}_2$ is the reversed version of $\mathbf{ord}_1$. The outputs of

the two modes are combined using a weighted linear combination, namely

$$\hat{\mathbf{s}}(n) = \Lambda\hat{\mathbf{s}}_{\mathbf{1}}(n) + (I - \Lambda)\hat{\mathbf{s}}_{\mathbf{2}}(n) \tag{5.3}$$

where Λ is an $M \times M$ diagonal weighting matrix, $\hat{\mathbf{s}}_{\mathbf{1}}(n)$ is the vector of symbol estimates from the normal mode MIMO DFE for the n-th time instant, $\hat{\mathbf{s}}_{\mathbf{2}}(n)$ is the vector of symbol estimates from the time-reversal mode MIMO DFE and $\hat{\mathbf{s}}(n)$ is the output of the linear combining block. Let $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_M$ be the diagonal entries of the matrix Λ. Then, the optimal choice for α_i is given by,

$$\alpha_i^{\text{opt}} = \frac{\sigma_{i,2}^2 - \rho_i\sigma_{i,1}\sigma_{i,2}}{\sigma_{i,1}^2 + \sigma_{i,2}^2 - 2\rho_i\sigma_{i,1}\sigma_{i,2}} \tag{5.4}$$

where $\sigma_{i,1}^2$, $\sigma_{i,2}^2$ and ρ_i are defined as in Section 2.5, but for user i.

The rationale for the proposed structure is as follows. While detecting the symbol $s_i(n)$, the normal mode MIMO DFE assumes that the interfering co-channel signals $\{s_1(n),\ s_2(n),\ \ldots,\ s_{i-1}(n)\}$ transmitted at the same time instant, n are known in addition to the "past" interfering symbols (both ISI and CCI). The time-reversal mode DFE, on the other hand, assumes knowledge of $\{s_M(n),\ s_{M-1}(n),\ \ldots,\ s_{i+1}(n)\}$ in addition to the "future" interfering symbols while detecting $s_i(n)$. In effect, all interfering symbols are assumed to be known and this will result in a performance improvement. Recall that, the knowledge of all the interfering symbols is a necessary condition to achieve the MFB.

5.4 Numerical Results

In this subsection, the MSE and SER performance improvements provided by the proposed MIMO LC-BiDFE structure over the conventional DFE structure will be demonstrated numerically[1]. Additional simulation examples are provided in Appendix A. We

[1]See [19] for Matlab source code

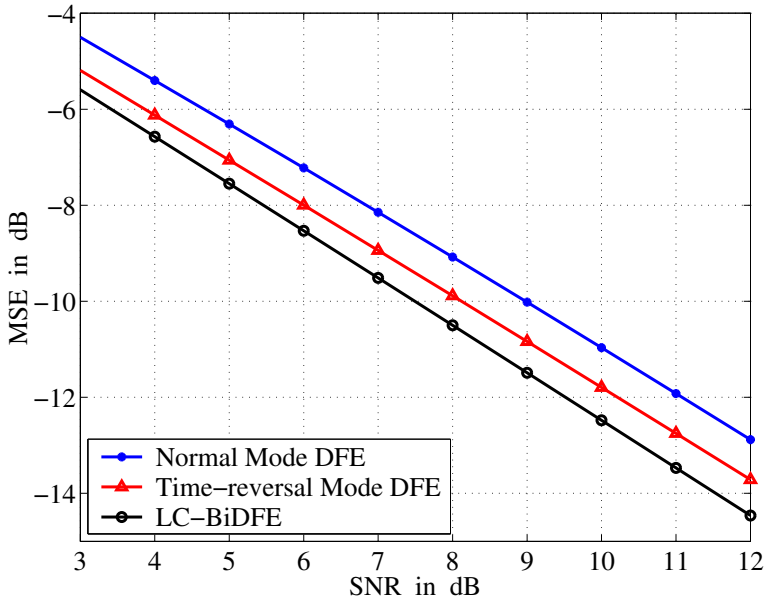

Figure 5.3: MSE performance of user 1 for the MIMO test channel C with $M = 2,\ P = 2$

consider a synthetic 2×2 MIMO channel with $L_c = 2$ taps. The MIMO impulse response of the channel C is given by,

$$\mathbf{c_{1,1}} = [0.9285,\ -0.3714],\ \mathbf{c_{1,2}} = [-0.6690,\ 0.7433],$$
$$\mathbf{c_{2,1}} = [0.9285,\ 0.3714],\ \mathbf{c_{2,2}} = [0.6690,\ 0.7433]. \tag{5.5}$$

The length of the FFF and FBF for the normal mode MIMO DFE were, $L_{f1} = 4$ and $L_{b1} = 1$, respectively. For the time-reversal mode MIMO DFE, a choice of $L_{f2} = 8$ and $L_{b2} = 1$ was made for the filter lengths. Unbiased MMSE coefficients were used for the MIMO DFEs for both the normal mode and time-reversal mode operation. The source symbols were selected from a BPSK source constellation. The order of detection for normal mode DFE was, $\mathbf{ord}_1 = [s_1(n),\ s_2(n)]$ while for the time-reversal mode DFE, $\mathbf{ord}_2 = [s_2(n),\ s_1(n)]$. The MSE performance curves for a normal mode DFE, time-reversal mode DFE and the LC-BiDFE are compared in Figures 5.3 and 5.4 for the two users $s_1(n)$ and $s_2(n)$, respectively. Ideal feedback was invoked to compute the MSE

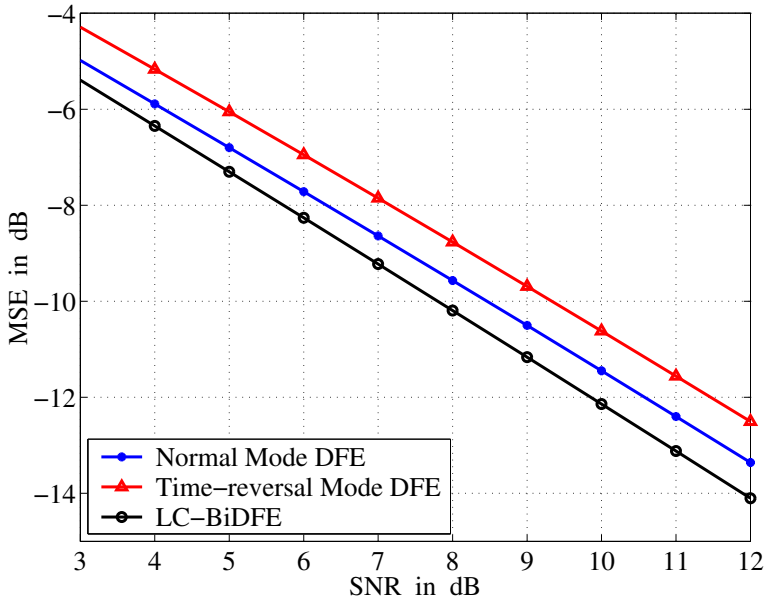

Figure 5.4: MSE performance of user 2 for the MIMO test channel C with $M = 2,\ P = 2$

values. In Figure 5.3, the time-reversal mode DFE has a lower MSE when compared to the normal mode DFE. This is due to the fact that, in the time-reversal mode DFE, user 1 is detected after user 2 and has additional knowledge of the interfering cursor symbol of user 2. The LC-BiDFE, however, has an additional gain of about 0.5 dB over the time-reversal mode DFE. Similarly, for user 2, the BiDFE has superior performance when compared to either the normal mode or the time-reversal mode DFEs and has a gain of at least 0.5 dB over the best performing constituent DFE (in this case, the normal mode DFE).

The symbol error rate (SER) curves for the normal mode DFE, time-reversal mode DFE and the LC-BiDFE are shown in Figures 5.5 and 5.6 for the two users $s_1(n)$ and $s_2(n)$, respectively, and the test channel C of equation (5.5). These simulation results do not invoke the assumption (2.5) and hence incorporate the effect of error propagation, i.e., decision feedback was assumed. For user 1, the soft MIMO LC-BiDFE offers a

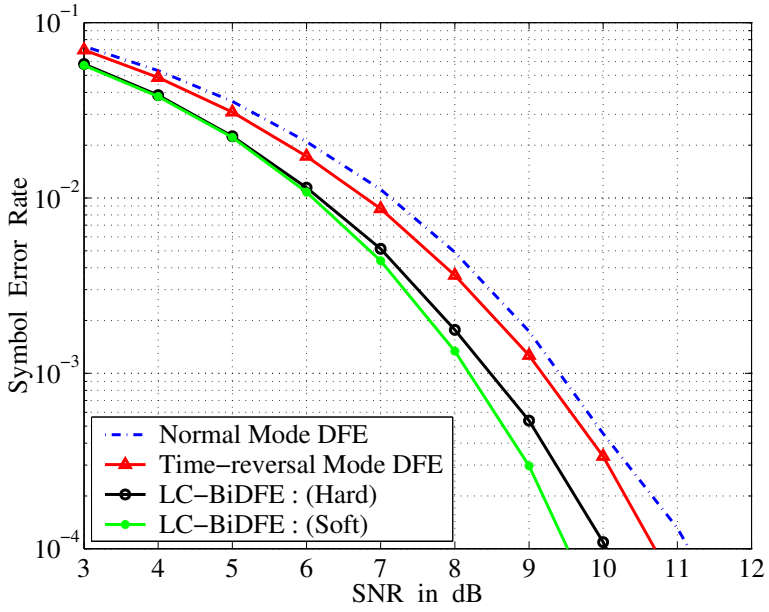

Figure 5.5: SER performance curves of user 1 for the MIMO test channel C with $M = 2,\ P = 2$

performance improvement of nearly 1 dB, at an SER of 10^{-3}, over the time-reversal mode MIMO DFE, while for user 2 the gain is about 0.6 dB for the soft MIMO LC-BiDFE over the normal mode MIMO DFE. These results demonstrate the viability of the LC-BiDFE in providing performance improvements in the MIMO channel equalization problem.

5.5 Summary

In this chapter, the idea of using time-reversal to improve the performance of a DFE has been extended to MIMO channel equalization. In addition, the MIMO DFE structure considered in this chapter affords the possibility of improved performance by rearranging the order of detection of the various cochannel users. Numerical results for a sample channel demonstrating achievable performance improvements has been provided. It has

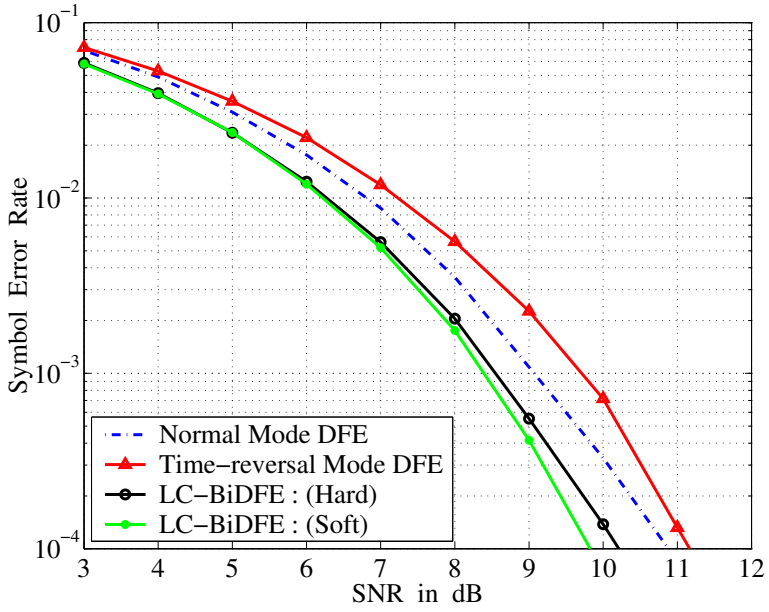

Figure 5.6: SER performance curves of user 2 for the MIMO test channel C with $M = 2,\ P = 2$

been further assumed that the MIMO channel impulse response is perfectly known at the receiver. In practice this condition is often violated and the channel impulse response needs to be estimated. Most packet based communication systems provide a training segment in each packet to facilitate channel estimation. During this phase, a sequence known to the receiver is transmitted. The design of such a training sequence for MIMO channel estimation will be considered in the next chapter.

Chapter 6

Training Sequence Design for MIMO Channel Estimation

In Chapter 5, the extension of a BiDFE to a multi-input multi-output (MIMO) channel equalization problem was considered. Knowledge of the channel impulse response from each transmit antenna to any receive antenna was assumed to be available at the receiver. However, in practice these channel impulse responses have to be estimated at the receiver. Training-based estimation, semi-blind estimation and blind estimation are three types of estimators that can be potentially used to estimate the MIMO channel impulse response at the receiver. Training-based estimators assume the presence of a training sequence and rely solely on the knowledge of these symbols. Blind estimation techniques, on the other hand, rely on qualitative information on the transmitted signals. Semi-blind techniques exploit the knowledge of the training symbols as well as the qualitative information about the data symbols.

A packet based system, for example GSM, typically contains a training sequence and hence training-based estimators are suitable. Although, semi-blind channel estimators may offer a better performance over purely training-based channel estimators [26],

purely training-based channel estimators are common in practice. The quality of the training-based channel estimate depends on the particular choice of training sequence. It has been known that training sequences with impulse-like auto-correlation functions are suitable for the single antenna channel estimation problem and the search for such sequences has received a great deal of attention in the past [68, 18]. The optimality criterion for training sequences for least-squares channel estimation has been considered in [24] and some optimal training sequences have been provided, but from the context of a single antenna system. On the other hand, near-optimal training sequences, derived using discrete Fourier transform techniques, are tabulated in [86].

The design of optimal training sequences for least-squares channel estimation in multiple antenna systems was considered and appropriate optimality criterion were derived by Balakrishnan et al. in [13, 89]. These results are presented in this chapter. In section 6.1, the system model is described and the least-squares channel estimator is reviewed. The optimality criterion for training sequence design is derived and the design trade offs in the choice of training sequence length are discussed in section 6.3. Section 6.4 describes a few heuristic methods for the search for near-optimal training sequences. A few near-optimal binary sequences obtained using these search methods, for employment in an EDGE (Enhanced Data for GSM Evolution) system, are tabulated in Appendix B. The special case of designing training sequences for a delay-diversity scheme is discussed in section 6.5 and section 6.6 concludes.

6.1 System Model

A system with M transmitter antennas and P receiver antennas is considered. Without loss of generality, each transmit antenna can be assumed to transmit different informa-

tion symbols. The received sampled signal vector at the p-th receive antenna during the n-th symbol period can be expressed by the discrete-time model,

$$r_p(n) = \sum_{m=1}^{M} \sum_{k=0}^{L-1} c_{m,p}(k) s_m(n-k) + w_p(n) \tag{6.1}$$

where $\mathbf{c}_{m,p} \triangleq [c_{m,p}(0), \ \ldots, \ c_{m,p}(L-1)]^T$ is the impulse response of the channel between the m-th transmit antenna and the p-th receive antenna, $s_m(n)$ is the transmitted symbol from the m-th transmit antenna and $w_p(n)$ is the additive noise at the p-th receive antenna. Each of the channel impulse responses is assumed to be of length L taps. During the training phase, different training sequences are transmitted from each of the transmit antennas. The training sequences are assumed to be N symbols long, and the channel impulse response is estimated at the receiver based on the knowledge of the training symbols. Here, the channel impulse responses $\mathbf{c}_{m,p}$ refer to the digital equivalent of the convolution of the transmit pulse shaping filter response with the physical multipath channel. Furthermore, the channel is assumed to be time-invariant for the duration of the packet.

6.2 MIMO Channel Estimation

Transmission of space-time coded signals requires coherent demodulation at the receiver. Hence, the receiver must be capable of estimating the channel from each transmit antenna. Current EDGE standards allow the transmission of a training sequence of length 26 symbols in each burst for a single transmit antenna system, which allows the receiver to estimate the channel for equalization. For estimating multiple channels from the multiple transmit antennas, we propose the use of different training sequences, one for each transmit antenna, transmitted simultaneously during each burst. These sequences have to be designed with good auto-correlation and cross-correlation properties

to enable accurate channel estimation at the receiver.

The vector of observations at the p-th receive antenna, during the training phase can be written in matrix form as,

$$\mathbf{r}_p = S\mathbf{c}_p + \mathbf{w}_p \tag{6.2}$$

where

$$\mathbf{r}_p \triangleq [r_p(N),\ \ldots,\ r_p(L)]^T \tag{6.3}$$

is the vector of observations,

$$\mathbf{c}_p \triangleq [c_{1,p}(0),\ \ldots,\ c_{1,p}(L-1),\ \ldots,\ c_{M,p}(0),\ \ldots,\ c_{M,p}(L-1)]^T \tag{6.4}$$

is the stacked vector of channel impulse responses,

$$\mathbf{w}_p \triangleq [w_p(N),\ \ldots,\ w_p(L)]^T \tag{6.5}$$

is the noise vector and S is an $(N-L+1) \times ML$ block-Toeplitz matrix consisting of the training symbols,

$$S \triangleq \begin{bmatrix} s_1(N) & \cdots & s_1(N-L+1) & \cdots & s_M(N) & \cdots & s_M(N-L+1) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ s_1(L) & \cdots & s_1(1) & \cdots & s_M(L) & \cdots & s_M(1) \end{bmatrix}. \tag{6.6}$$

The stacked impulse response vector $\mathbf{c}_p$ is estimated at the receiver for each of the P receiver antennas. A least-squares (LS) channel estimator is considered for this purpose. An LS channel estimator minimizes the squared error between the received signal vector and the channel estimate-based reconstructed signal. The LS estimate is given by,

$$\begin{aligned} \hat{\mathbf{c}}_p^{\text{LS}} &\triangleq \arg\min_{\mathbf{c}} \|\mathbf{r}_p - S\mathbf{c}\|^2 \\ &= \left(S^H S\right)^{-1} S^H \mathbf{r}_p. \end{aligned} \tag{6.7}$$

In equation (6.7) the auto-correlation matrix $S^H S$ is assumed to be invertible. The LS estimate can be expressed in terms of the channel impulse response as,

$$\hat{\mathbf{c}}_p^{\mathrm{LS}} = \mathbf{c}_p + \left(S^H S\right)^{-1} S^H \mathbf{w}_p. \tag{6.8}$$

If the additive noise is zero-mean and uncorrelated to the training sequence, the LS channel estimate is unbiased. The LS channel estimation error is,

$$\mathrm{E}\left[\|\mathbf{c}_p - \hat{\mathbf{c}}_p^{\mathrm{LS}}\|^2\right] = \mathrm{tr}\left\{\left(S^H S\right)^{-1} S^H \mathrm{E}\left[\mathbf{w}_p \mathbf{w}_p^H\right] S \left(S^H S\right)^{-1}\right\}. \tag{6.9}$$

Under the assumption that the noise process is white with a variance of σ_w^2, the LS estimation error simplifies to

$$\mathrm{E}\left[\|\mathbf{c}_p - \hat{\mathbf{c}}_p^{\mathrm{LS}}\|^2\right] = \mathrm{tr}\left\{\sigma_w^2 \left(S^H S\right)^{-1}\right\}. \tag{6.10}$$

6.3 Training Sequence Design

From equation (6.10), it is evident that the LS estimation error depends on the choice of training sequence. Hence, the training sequence S can be optimized to minimize the LS estimation error.

Theorem 6.1 *The minimum LS error is obtained if and only if*

$$S^H S = (N - L + 1)\sigma_t^2 I_{ML}, \tag{6.11}$$

where σ_t^2 is the maximum permissible variance for the training symbols.

Proof: The arguments outlined here are similar to those provided in [85]. The training sequence is optimized to minimize the LS channel estimation error of equation (6.10). This is equivalent to minimizing $\mathrm{tr}\left\{\left(S^H S\right)^{-1}\right\}$ with respect to the training sequence S. Let $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_{ML}$ be the eigenvalues of the auto-correlation matrix $S^H S$. Since the

Hermitian matrix $S^H S$ is assumed to be invertible, all the eigenvalues are greater than zero. Let us now define the vectors

$$\underline{\lambda} = \left[\sqrt{\lambda_1},\ \ldots,\ \sqrt{\lambda_{ML}}\right]^T; \quad \underline{\lambda}^{'} = \left[\frac{1}{\sqrt{\lambda_1}},\ \ldots,\ \frac{1}{\sqrt{\lambda_{ML}}}\right]^T. \tag{6.12}$$

By the Cauchy-Schwartz inequality,

$$\langle \underline{\lambda}, \underline{\lambda}^{'} \rangle^2 \leq \langle \underline{\lambda}, \underline{\lambda} \rangle \langle \underline{\lambda}^{'}, \underline{\lambda}^{'} \rangle \tag{6.13}$$

where $\langle , \rangle$ denotes a vector dot product. Hence,

$$(ML)^2 \leq \left(\sum_{k=1}^{ML} \lambda_k\right) \left(\sum_{k=1}^{ML} \frac{1}{\lambda_k}\right). \tag{6.14}$$

Since, the sum of eigenvalues is the trace of a matrix, the above inequality can be rewritten as,

$$\text{tr}\left\{\left(S^H S\right)^{-1}\right\} \geq \frac{(ML)^2}{\text{tr}\left\{S^H S\right\}}. \tag{6.15}$$

Equality occurs in equation (6.15) only when $S^H S = \alpha I$ for some constant α, i.e, $S^H S$ must be proportional to the identity matrix. Further more, the tr$\{S^H S\}$ can be maximized by choosing the maximum energy symbols from the source constellation, namely the farthest points from the origin. Let σ_t^2 be the variance of these maximum energy source symbols chosen for training. Then, the minimum value of the estimation error is

$$\min \text{E}\left[\|\mathbf{c}_p - \hat{\mathbf{c}}_p^{\text{LS}}\|^2\right] = \frac{ML\sigma_w^2}{(N-L+1)\sigma_t^2}. \tag{6.16}$$

The above result is analogous to the single transmit antenna scenario and is equivalent to choosing the training sequences to be temporally white and spatially uncorrelated (i.e., across transmit antennas). The LS channel estimator turns out to be the best (in terms of having the minimum mean-squared error) among all unbiased estimators and it is the most efficient in the sense that it achieves the Cramer-Rao lower bound [51].

Hence, the use of an LS estimator is considered to derive the optimality criterion for training sequence design. Although, the optimality condition for the training sequences have been derived assuming that the channel is time-invariant during the duration of the slot, they would still be valid if the channel is slowly time varying. If the channel time variations are rapid, channel tracking is essential and purely training based channel estimators are no longer suitable.

6.3.1 Training Sequence Length Design

A critical parameter in training sequence design is the length of the sequence. The training sequence needs to be long enough for the channel to be identified. A longer training sequence has the added advantage of reducing the channel estimation error. However, an increase in training sequence length results in a decrease in the useful data rate of the transmission. The design trade offs associated with the choice of training sequence length are discussed in this subsection.

Identifiability

For the channel impulse response to be identifiable, the auto-correlation matrix $S^H S$ of equation (6.7) has to be invertible. Hence, the training sequence matrix S has to be of full column rank. The necessary condition for S to be full column rank is,

$$(N - L + 1) \geq ML. \tag{6.17}$$

Loss due to Channel Estimation

Any error resulting from channel estimation can be incorporated into the noise process and can be quantified as a loss in effective SNR. During data transmission we have

(recall equation (6.1)),

$$r_p(n) = \sum_{m=1}^{M}\sum_{k=0}^{L-1} c_{m,p}(k)s_m(n-k) + w_p(n), \tag{6.18}$$
$$= \sum_{m=1}^{M}\sum_{k=0}^{L-1} c^{\text{LS}}_{m,p}(k)s_m(n-k) + \underbrace{\sum_{m=1}^{M}\sum_{k=0}^{L-1}\left\{c_{m,p}(n) - c^{\text{LS}}_{m,p}(k)\right\} s_m(n-k) + w_p(n)}_{w^e_p(n)}$$

where the term $w^e_p(n)$ denotes the equivalent noise. However, the statistics of the noise process is no longer Gaussian. However, simulations (see [89]) show that it is possible to treat the noise term as Gaussian with an equivalent noise variance. Let us assume that the equivalent noise $w^e_p(n)$ is uncorrelated with the source symbols and the source symbols to be i.i.d with a variance σ_s^2. The equivalent noise variance (or the MSE) is then given by,

$$\begin{aligned}\text{MSE} &= \text{E}\left[|w^e_p(n)|^2\right] \\ &\stackrel{a}{=} \text{E}\|\mathbf{c}_p - \hat{\mathbf{c}}^{\text{LS}}_p\|^2\sigma_s^2 + \sigma_w^2 \\ &\stackrel{b}{=} \sigma_w^2\left\{1 + \frac{ML\sigma_s^2}{(N-L+1)\sigma_t^2}\right\}.\end{aligned} \tag{6.19}$$

In equation (6.19), $\stackrel{a}{=}$ is obtained by using the i.i.d assumption on the source symbols and the vector representation of the stacked channel impulse response, while step $\stackrel{b}{=}$ results from equation (6.16). If the source symbols were to be chosen from a constant modulus type constellation (for example 8-PSK), then the source variance will be the same as the variance of the training symbols and equation (6.19) can be further simplified. The increase in MSE due to the channel estimation error can be interpreted as a loss in effective SNR at the receiver.

Loss in Throughput

The throughput of a system is the product of the data rate and the probability of successful transmission of a packet. Clearly, from equation (6.19), the longer the training sequence, the lesser the loss in effective SNR. A smaller channel estimation error would result in a decrease in packet error rate. However, an increase in the training sequence length reduces the number of information bits that can be transmitted in a packet and hence the data rate. An EDGE packet consists of 26 training symbols and 118 data symbols. Hence, any additional training symbols come at the cost of the data symbols. It is clear from the above discussion that the throughput depends on the length of the training sequence. Hence, a good criterion for the design of training length would be to maximize the achievable throughput. The achievable throughput for a given choice of training length can be calculated if the probability distribution function (PDF) of the SINR, the achievable data rate and the corresponding block error rate are known. The average throughput is then calculated as,

$$\text{Average Throughput} = \int_{\infty}^{\infty} f(\gamma)\left\{1 - P_e(\gamma, N, L)\right\} R(N) d\gamma \tag{6.20}$$

where $f(\gamma)$ is the PDF of the SINR, $P_e(\gamma, N)$ and $R(N)$ are the block error rates and the data rates, respectively, as a function of the SINR, the training length N and the channel impulse response length parameter L.

6.4 Search for good Training Sequences

Once the training sequence length has been decided, it becomes essential to search for training sequences with good properties. Equation (6.11) specifies the optimality criterion for the search of such sequences. The sequences have to be of simple alphabets in order to guarantee low complexity realization. Two types of sequences are common:

aperiodic and periodic sequences. While aperiodic sequences exist for many lengths, periodic ones are much harder to find. However, due to an expansion theorem [79], short periodic sequences can be concatenated to very large sequences preserving their orthogonal properties. A list of known periodic sequences of simple alphabets are tabulated in [13]. Originally applied to single antenna systems, the periodic sequences can be used for multiple antenna systems. Construction of multiple antenna training sequences from the periodic sequence is described later in subsection 6.4.3. The QPSK sequence of length $L = 16$ was proposed to extend existing OFDM systems to four transmit and receive antennas [71].

It is possible that optimum training sequences may not exist for a particular choice of training length and channel delay spread. In that case, training sequences with near optimal properties can be searched for. A few heuristic methods for the search of such sequences are discussed in this section. A few near-optimal binary training sequences, suitable for a multi-antenna EDGE system, are listed in Appendix B. The search for these sequences are based on the methods described in this section.

6.4.1 Full search

Near-optimal sequences can be obtained by searching over all possible sequences and choosing those which have the minimum value of $\mathrm{tr}\{(S^H S)^{-1}\}$. However, the search has to be done over $|\mathcal{A}|^{MN}$ sequences, where $|\mathcal{A}|$ is the number of points in the source constellation. This search is computationally prohibitive. Hence, heuristic methods that search over a reduced set of sequences are of special interest.

6.4.2 Random Search

From equation (6.11) it is clear that near-optimal sequences should have good auto-correlation and cross-correlation properties, i.e., small non-peak auto-correlation terms over a window of size $L-1$ on either side of the peak location and small cross-correlation terms for a window of length $2L-1$. To begin with, sequences with good auto-correlation properties can be determined by searching over all the $|\mathcal{A}|^N$ possible sequences. The number of such sequences can be expected to be much smaller than $|\mathcal{A}|^{MN}$. This is followed by a search for M sequences with good cross-correlation properties from this reduced set of sequences.

6.4.3 Cyclic Shift Search

Consider the sequence $\mathbf{t}_1 = [s(1),\ \ldots,\ s(N')]$ of length N', where $N' = N - L + 1$. The sequences $\{\mathbf{t}_2,\ \ldots,\ \mathbf{t}_M\}$ are now constructed by cyclic-shifts of the sequence $\mathbf{t}_1$. For example, the sequence $\mathbf{t}_{k+1} = [s(k\delta+1),\ \ldots,\ s(N'),\ s(1),\ \ldots,\ s(k\delta)]$ is obtained by a cyclic-shift of $k\delta$ of the sequence $\mathbf{t}_1$, where $\delta = \lfloor \frac{N'}{M} \rfloor$. New sequences $\{\mathbf{s}_1,\ \ldots,\ \mathbf{s}_M\}$ are constructed by adding a cyclic-prefix of length $L-1$ to the sequences $\{\mathbf{t}_1,\ \cdots,\ \mathbf{t}_M\}$. For example, $\mathbf{s}_1 = [s(N'-L+2),\ \ldots,\ s(N'),\ s(1),\ \ldots,\ s(N')]$ is one such sequence derived from the original sequence. Note that the new sequences $\mathbf{s}_k$ are of length N.

If the sequence $\mathbf{t}_1$ has a cyclic auto-correlation function with zero off-peak terms and if $\delta \geq L$, then it is easy to see that equation (6.11) will be satisfied for the choice of training sequences $\mathbf{s}_1,\ \ldots,\ \mathbf{s}_M$. However, when searching for near-optimal training sequences the restriction of zero off-peak terms for the cyclic auto-correlation function can be relaxed and small off-peak cyclic auto-correlation terms can be allowed. This restricts the search space to a size $|\mathcal{A}|^{N-L+1}$.

6.5 Training Sequence for Delay-diversity Scheme

A delay-diversity scheme is a simple space-time coding scheme that is used to achieve diversity when multiple transmit antennas are available. We consider a two transmit antenna system employing delay diversity. In this scheme, the same information symbols are transmitted from two antennas with a single symbol delay on the second antenna. Thus, the information symbols are received twice with different path gains resulting in diversity gain. The delay diversity technique can also be viewed as a trellis space-time code [84] and has the advantage that an optimized equalizer is sufficient to decode the delay-diversity space-time code.

For the delay-diversity scheme, the received signal at the p-th receive antenna for the n-th symbol period is,

$$\begin{aligned} r_p(n) &= \sum_{m=1}^{2} \sum_{k=0}^{L-1} c_{m,p}(k) s(n-k-m+1) + w_p(n) \\ &= \sum_{k=0}^{L} \{\tilde{c}_{1,p}(k) + \tilde{c}_{2,p}(k)\} s(n-k) + w_p(n) \end{aligned} \tag{6.21}$$

where $\tilde{\mathbf{c}}_{1,p}$ and $\tilde{\mathbf{c}}_{2,p}$ are augmented channel impulse response vectors of length $L+1$ taps such that $\tilde{\mathbf{c}}_{1,p} = [c_{1,p}(0),\ \ldots,\ c_{1,p}(L-1),\ 0]^T$ and $\tilde{\mathbf{c}}_{2,p} = [0,\ c_{2,p}(0),\ \ldots,\ c_{2,p}(L-1)]^T$. The equivalent channel impulse response $\mathbf{c}_p^{\text{eq}}$ is a sum of these augmented channel impulse response vectors and it is sufficient to estimate $\mathbf{c}_p^{\text{eq}}$ at the receiver.

During the training phase, the transmitter, however, has the option of transmitting different training sequences from each of the two transmitter antennas. The individual channel impulse responses can then be estimated and summed up with the appropriate delay to obtain the equivalent channel impulse response. A second possibility would be to transmit the same training sequence from both the transmitter antennas, one of them with a single symbol delay, and estimate the equivalent channel response directly. Intuitively, the idea of using the same training sequence is appealing and it will be shown

that this, in fact, is a better choice.

First, we consider the use of different training sequences from the two transmit antennas. For the framing structure to be preserved in a delay diversity scheme, it is essential that the training sequence be delayed in a fashion similar to the data. Hence, the training sequence from the second antenna is transmitted with a symbol delay. The received sequence at the p-th receive antenna, during the training phase, can be expressed as,

$$\begin{aligned} r_p(n) &= \sum_{m=1}^{2} \sum_{k=0}^{L-1} c_{m,p}(k) s_m(n-k-m+1) + w_p(n) \\ &= \sum_{k=0}^{L} \{\tilde{c}_{1,p}(k) s_1(n-k) + \tilde{c}_{2,p} s_2(n-k)\} + w_p(n) \end{aligned} \tag{6.22}$$

and the stacked channel impulse response vector $\tilde{\mathbf{c}}_p$ can be estimated as in equation (6.7). The only difference is that $\tilde{\mathbf{c}}_p$ has $2(L+1)$ taps and S is an $(N-L) \times 2(L+1)$ block-Toeplitz matrix. The equivalent channel impulse response $\mathbf{c}_p^{\text{eq}}$ can be expressed in terms of $\tilde{\mathbf{c}}_p$ as,

$$\mathbf{c}_p^{\text{eq}} = A \tilde{\mathbf{c}}_p, \tag{6.23}$$

where the $(L+1) \times 2(L+1)$ matrix A is constructed as,

$$A = \begin{bmatrix} I_L & 0_{L\times 2} & 0_{1\times L} \\ 0_{1\times L} & 0_{1\times 2} & I_L \end{bmatrix}. \tag{6.24}$$

The LS error for the equivalent channel impulse response is,

$$\mathrm{E}\left[\|\mathbf{c}_p^{\text{eq}} - \hat{\mathbf{c}}_p^{\text{eq}}\|^2\right] = \mathrm{tr}\left\{\sigma_w^2 A \left(S^H S\right)^{-1} A^T\right\}. \tag{6.25}$$

The training sequence S has to be optimized to minimize the LS channel estimation error of equation (6.25).

Theorem 6.2 *The least possible LS channel estimation error, when different training sequences are used from the two transmit antennas in a delay-diversity scheme, is given by*

$$E\left[\|\mathbf{c}_p^{eq} - \hat{\mathbf{c}}_p^{eq}\|^2\right] \geq \frac{\sigma_w^2(L+1)}{\sigma_t^2(N-L)}. \tag{6.26}$$

Proof: The LS estimation error of the equivalent channel impulse response is expressed in equation (6.25). The training sequence S is to be optimized to minimize this estimation error. Let $R = S^H S$ denote the auto-correlation matrix of the training sequence S. The matrix R is Hermitian and positive definite. We assume that the maximum energy points of the source constellation are used as the training symbols. Hence the diagonal entries of R are equal to $\sigma_t^2(N-L)$.

Let $\mathbf{a}_1, \ \ldots, \ \mathbf{a}_{L+1}$ be the row vectors of the augmentation matrix A. Each of this row vectors $\mathbf{a}_k$ has non-zero entries, namely unity, at the k-th location and the $(L+1+k)$-th location, except for $\mathbf{a}_1$ and $\mathbf{a}_{L+1}$. The vectors $\mathbf{a}_1$ and $\mathbf{a}_{L+1}$ are unit vectors with an entry of one at the first and the $2(L+1)$-th locations, respectively. The LS error of equation (6.25) can be expressed as a function of the row vectors $\mathbf{a}_k$ as,

$$\mathrm{E}\left[\|\mathbf{c}_p^{\mathrm{eq}} - \hat{\mathbf{c}}_p^{\mathrm{eq}}\|^2\right] = \sigma_w^2 \sum_{k=1}^{L+1} \left\langle R^{-1}\mathbf{a}_k, \mathbf{a}_k \right\rangle . \tag{6.27}$$

Kantorovich inequality: If $R \in \mathcal{C}^{n\times n}$ is a positive definite Hermitian matrix and $x \in \mathcal{C}^n$ is a vector, then

$$\langle x, x\rangle^2 \leq \langle Rx, x\rangle\langle R^{-1}x, x\rangle. \tag{6.28}$$

Kantorovich inequality [63] can be applied to equation (6.27) and a lower bound for the LS estimation error can be obtained. For a k value of $2, \ \ldots, \ L$ one can see that $\left\langle \mathbf{a}_k, \mathbf{a}_k \right\rangle = 2$ and

$$\left\langle R\mathbf{a}_k, \mathbf{a}_k \right\rangle = r_{k,k} + r_{k,k+L+1} + r_{k+L+1,k} + r_{k+L+1,k+L+1}, \tag{6.29}$$

where $r_{i,j}$ are the entries of the auto-correlation matrix R. Auto-correlation matrices have the property that the off-diagonal entries are no larger than the diagonal entries. Hence,

$$\langle R\mathbf{a}_k, \mathbf{a}_k \rangle \leq 4\sigma_t^2(N-L). \tag{6.30}$$

From equations (6.28) and (6.30), it is clear that for the particular choice of k between 2 to L,

$$\langle R^{-1}\mathbf{a}_k, \mathbf{a}_k \rangle \geq \frac{1}{\sigma_t^2(N-L)}. \tag{6.31}$$

Equation (6.31) is also satisfied when $k = 1$ and $k = L+1$. Hence equation (6.27) reduces to,

$$\mathrm{E}\left[\|\mathbf{c}_p^{\mathrm{eq}} - \hat{\mathbf{c}}_p^{\mathrm{eq}}\|^2\right] \geq \frac{\sigma_w^2(L+1)}{\sigma_t^2(N-L)}. \tag{6.32}$$

■

Now, consider the use of the same training sequence from both the transmit antennas, with the appropriate symbol delay. From equation (6.21), the LS estimate of the equivalent channel impulse response $\mathbf{c}_p^{eq}$ can be determined like that of an SISO system. Since, the same training sequence is used from both the antennas, the Toeplitz matrix S has a size $(N-L)\times(L+1)$. Based on analysis similar to Theorem 6.1, it can be shown that the minimum possible LS estimation error is,

$$\min \mathrm{E}\left[\|\mathbf{c}_p^{eq} - \hat{\mathbf{c}}_p^{eq}\|^2\right] = \frac{\sigma_w^2(L+1)}{\sigma_t^2(N-L)}, \tag{6.33}$$

and is obtained if and only if

$$S^H S = (N-L)\sigma_t^2 I_{L+1}. \tag{6.34}$$

Comparing equation (6.33) with the lower bound obtained in equation (6.26), it is clear that the choice of identical training sequences for the two transmit antennas, provided equation (6.34) is satisfied, is indeed optimal. The above result is not restricted to

the case of a single symbol delay in the delay-diversity scheme. The optimality of the choice of identical training sequences is also valid for a delay-diversity scheme with a symbol delay δ. In such a scenario, the A matrix will be of size $(L+\delta) \times 2(L+\delta)$, c_p^{eq} will have $2(L+\delta)$ taps and S will be an $(N-L-\delta+1) \times 2(L+\delta)$ block-Toeplitz matrix. Using similar arguments as before it can be shown that

$$\mathrm{E}\left[\|\mathbf{c}_p^{eq} - \hat{\mathbf{c}}_p^{eq}\|^2\right] \geq \frac{\sigma_w^2(L+\delta)}{\sigma_t^2(N-L-\delta+1)}, \tag{6.35}$$

and that the lower bound can be achieved if identical training sequences (satisfying a condition analogous to equation (6.34)) are used.

6.6 Summary

In this chapter, the channel impulse response for multiple antenna systems was determined using least squares channel estimation. It was shown that the loss in throughput due to channel estimation can be minimized by appropriate choice of training length and training sequences. These optimal sequences ought to satisfy the property of both temporal and spatial whiteness, i.e., have a low auto-correlation as well as low cross-correlation. After deriving an optimality criterion for training sequence design for MIMO systems, a few heuristic methods for the search of near-optimal sequences were proposed. It was further shown that the choice of identical training sequences, transmitted with the appropriate delays, is optimal for a two transmit antenna system employing the delay-diversity scheme. When identical training sequences are transmitted from the two antennas, the estimation of the equivalent channel impulse response of equation (6.23), at the receiver, is no different from that of estimating the channel impulse response for a 1x1 antenna configuration. This implies that the receiver architecture needs no modification for a 2x1 antenna configuration, employing a delay diversity scheme,

except the capability to estimate and equalize a longer effective channel impulse response when compared to the 1x1 antenna configuration.

Chapter 7

Conclusions

"One always has time enough, if one will apply it well."

- Goethe

Equalization approaches have traditionally relied on causal processing of the signal at the receiver. The increase in popularity of packet based transmission systems like the GSM or EDGE offers the possibility of block processing of the received signal. With block processing comes the freedom to process the signal in either a causal or a non-causal fashion. The advantages of employing time-reversal has been presented in this dissertation from the context of a decision feedback equalizer. A summary of the results in this dissertation are presented in Section 7.1 and future research directions are discussed briefly in Section 7.2.

7.1 Summary of Results

Firstly, the performance and limitations of a DFE as an equalizer structure were reviewed. A novel bidirectional DFE architecture that employs time-reversal of the received block of data was proposed. A BiDFE consists of two parallel DFE structures,

one to equalize the received signal and the other the time-reversed version of the received signal. The causal nature of error propagation causes the error bursts to proceed in opposite directions in the two parallel DFEs and results in a low correlation between the error bursts. Error propagation can be mitigated by combining the outputs of the two parallel DFEs. In addition, under the ideal feedback assumption, the BiDFE, effectively, assumes knowledge of both the past and future interfering symbols and hence decreases the gap from the MFB. The BiDFE architecture (BAD and LC-BiDFE) was shown to provide a significant performance improvement over a conventional DFE with little additional complexity. One disadvantage of the BiDFE is the increase in latency or the overall detection delay.

In an attempt to gain insight on the performance limitation of the LC-BiDFE, the finite length constraint was relaxed and the asymptotic (as the noise variance approaches zero) MSE performance was evaluated, under the ideal feedback assumption. Although, the infinite length LC-BiDFE offers performance improvement over the conventional DFE, the gap from the MFB was non-zero. This led to the formulation of the tap optimization problem, where the coefficients of the LC-BiDFE were optimized to minimize the overall MSE instead of the MSE at the output of each of the constituent DFEs. The MMSE-BiDFE coefficients were determined and were shown to attain the MFB. It was further shown by a numerical example that the LC-BiDFE tap setting that attains the MFB is not unique.

The taps of the LC-BiDFE were optimized under a finite length constraint for the special case of a symmetric channel. The effectiveness of the tap optimization in an LC-BiDFE was evaluated numerically. It was shown that, although the MMSE-BiDFE offered an improved performance under the ideal feedback assumption, there was no improvement in SER when decision feedback was used. The MSE cost function was mod-

ified by incorporating the norm of the feedback filter tap coefficients and the LC-BiDFE taps that minimizes the modified cost function was shown to offer an SER improvement, albeit marginal. An alternate solution, based on the use of an iterative BiDFE approach, was demonstrated to be reasonably successful in offering an SER improvement from tap optimization, when the ideal feedback assumption was violated. The tap optimization was also extended to the asymmetric channel scenario, by constraining the receiver structure to have a front-end filter, matched to the channel impulse response.

The BiDFE was extended to multiple-input multiple-output (MIMO) channel equalization. The effectiveness of a MIMO LC-BiDFE was demonstrated with simulations. One aspect in which the MIMO BiDFE differs from the SISO BiDFE is in the use of different order of detection for the users in the two parallel MIMO DFE streams.

In training based MIMO channel estimation the choice of training sequence and length affects performance. Training sequences that satisfy the property of being white across both time and space (i.e., antennas) were shown to be optimal. For a delay diversity scheme, transmitting the same optimal (in the single user sense) training sequence from both transmit antennas was shown to be optimal for LS channel estimation. Hence, the receiver architecture requires no change from the single antenna case, when multiple transmit antennas are deployed and a delay diversity scheme is used.

7.2 Future Directions

In Chapter 3, an MMSE-BiDFE tap setting that attains the MFB was derived. It was shown, with the aid of a 2-tap channel example, that the LC-BiDFE filter settings that attain the MFB are not unique. It will be interesting to know, if the lack of uniqueness of MFB attaining LC-BiDFE filter settings extend to all generic channel impulse responses.

Furthermore, knowledge of all such solutions would help in choosing the "best" filter setting, when the ideal feedback assumption is relaxed.

In Chapter 4, the tap optimization of the LC-BiDFE resulted in an MSE performance improvement, under the ideal feedback assumption. However, this performance improvement did not translate into a decrease in the SER, when the ideal feedback assumption was relaxed. Two solutions, one based on the optimization of the modified cost function and the other using an iterative LC-BiDFE approach, proposed to mitigate this problem offered only a marginal improvement. Furthermore, this marginal improvement could be attained only if the user defined parameter ξ (for the modified cost function) is chosen appropriately. In the case of the iterative LC-BiDFE approach, the performance improvement comes at the cost of increased computational complexity. The search for a low-complexity solution that can result in the realization of the performance improvement due to tap optimization, in the presence of decision feedback, needs further investigation.

In Chapter 5, the BiDFE architecture has been extended to the MIMO channel equalization problem. There are a number of open problems in this area, including performance analysis of the MIMO LC-BiDFE and tap optimization of the MIMO LC-BiDFE. Furthermore, there is no known extension of the bidirectional arbitrated decision feedback equalizer (BAD) scheme to the equalization of a MIMO channel with memory.

Appendix A

Additional Simulation Results

In this appendix, additional simulation examples demonstrating the performance improvement capabilities of the BiDFE are presented. Matlab script files for generating the simulation plots provided in this appendix are available at [19]. Frequency-selective fading channel models for a typical urban environment are considered in Section A.1 to supplement the simulation examples of Subsection 2.5.6. Section A.2 considers a synthetic channel impulse response with severe ISI, namely roots close to the unit circle, and compares the performance of the LC-BiDFE with that of BAD and MLSE. Section A.3 provides an additional simulation example to demonstrate the efficacy of the iterative LC-BiDFE approach proposed in Section 4.3. Additional fading channel simulations to test the efficacy of the MIMO LC-BiDFE are provided in Section A.4.

A.1 Broadband Wireless Channels

In this section, the equalization of quasi-static frequency-selective fading channel impulse responses is considered. Each realization of the channel impulse response (corresponding to the channel encountered by each packet of data) is modeled based on the

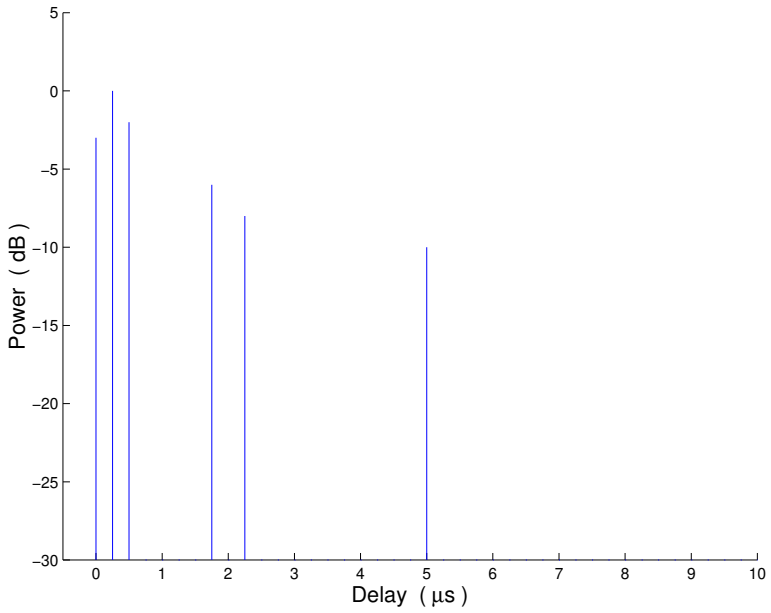

Figure A.1: Power profile of fading channel for a typical urban environment

multipath delay profile specified by GSM [32] (with slight modifications) for the urban environment. The power profile for the urban environment is shown in Figure A.1. All the paths in the above profile were assumed to be independently Rayleigh fading. A square-root raised cosine (SRRC) pulse shaping filter with a 12.5% roll-off factor was used at the transmitter. A symbol rate of 1 Mbaud (the symbol period is 1 μs) was assumed and the source was drawn from a BPSK constellation. A total of 16 feedforward taps and 8 feedback taps were allocated to the DFE. Perfect channel knowledge (assumption 2.4) was assumed at the receiver.

Simulations were performed for an ensemble of multipath channels, randomly generated using the urban environment power profile. In Figure A.2, the SER performance (averaged over the various channel realizations) of an LC-BiDFE (with hard/soft decisions) is compared with the performance of a normal mode DFE and BAD. For the fading channel simulations, BAD offers a performance improvement of 1.1 dB, at an

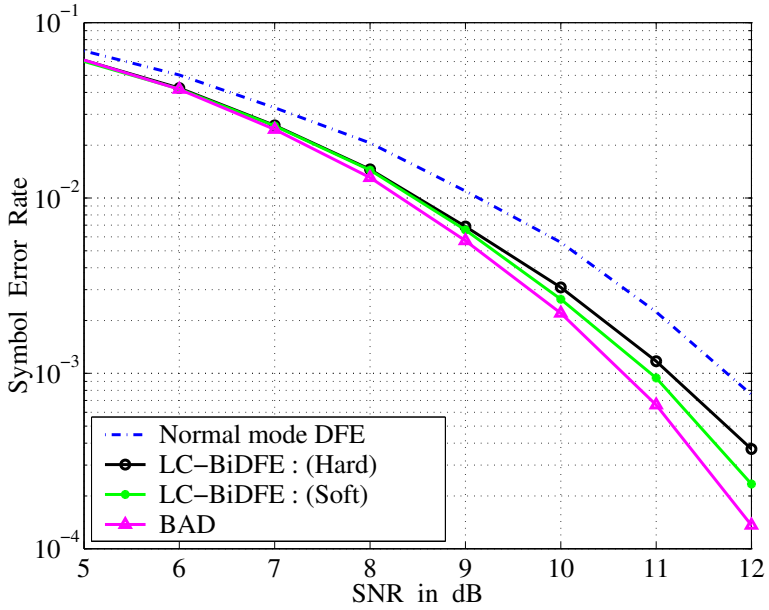

Figure A.2: SER performance comparison for a fading channel environment

SER of 10^{-3}, over the normal mode DFE, while the LC-BiDFE (with soft decisions) offers about 0.8 dB of improvement. Although, BAD offers better improvement than LC-BiDFE, the improvement is small and comes at the cost of increased receiver complexity. The SER performance curves for four randomly generated (sample) channels with the urban power profile is illustrated in Figure A.3.

A.2 BiDFE Performance for a Channel with Deep Nulls

In this section we consider a synthetic channel impulse response with a deep null. Such a channel will have roots close to the unit circle. The root locations of the channel,

$$\mathbf{c}_3 = [-0.3920 \;\; 0.0152 \;\; 0.8320 \;\; 0.0152 \;\; -0.3920] \tag{A.1}$$

are illustrated in Figure A.4. The SER performance of a normal mode DFE is compared with that of an LC-BiDFE, BAD and MLSE in Figure A.5. The source is assumed to

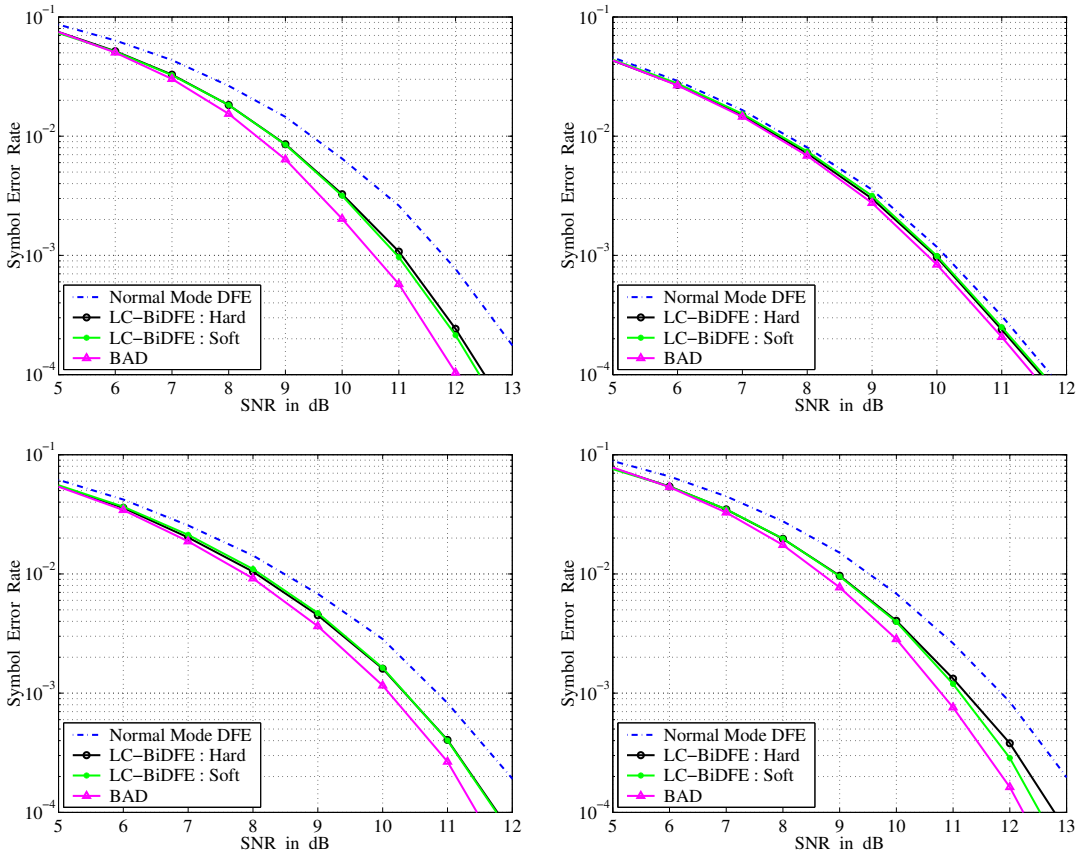

Figure A.3: SER performance comparison for a few sample channels with the urban power profile

be drawn from a BPSK source. A DFE with 8 feedforward filter taps and 4 feedback filter taps was used. The LC-BiDFE offers only a small gain of about 0.4 dB from the normal mode DFE curve at an SER value of 10^{-3}, while BAD offers a gain of nearly 1.1 dB. When the channel has deep nulls, the magnitudes of the FBF taps tends to be large and the error propagation phenomenon is more severe. As the reconstruction-and-arbitration scheme of BAD is better designed towards exploiting the low correlation in the error bursts between the normal mode DFE and time-reversal mode DFE streams, it offers a significantly better performance. However, the gap from the MLSE for the BAD algorithm is about 0.8 dB.

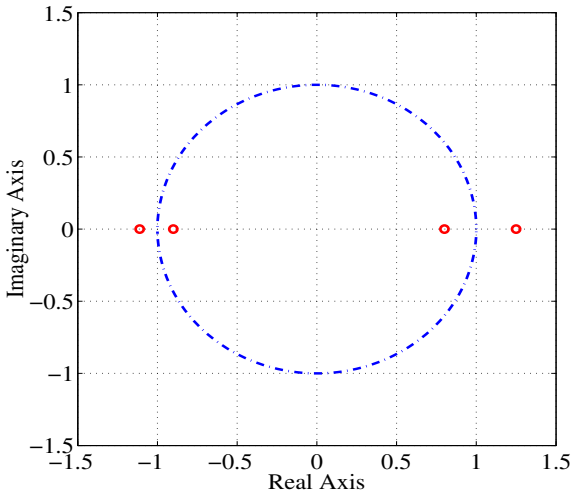

Figure A.4: Channel zeros for $\mathbf{c}_3 = [-0.3920 \;\; 0.0152 \;\; 0.8320 \;\; 0.0152 \;\; -0.3920]$

A.3 Simulation Example for LC-BiDFE with Iteration

In Section 4.3, an iterative LC-BiDFE approach was proposed to ensure that the performance improvement obtained with tap optimization is attained, at least partly, when the ideal feedback assumption is relaxed. In this section, an additional simulation example to supplement the numerical result in Section 4.3 is provided. We consider the symmetric channel impulse response,

$$\mathbf{c}_4 = [-0.2053 \;\; 0.1300 \;\; 0.9391 \;\; 0.1300 \;\; -0.2053]. \tag{A.2}$$

Recall that, in the iterative LC-BiDFE architecture, proposed in Section 4.3, the received signal is first equalized by using an LC-BiDFE with MMSE-DFE settings. This is followed up by equalizing the received signal using an LC-BiDFE with MMSE-BiDFE tap settings. The SER performance of the iterated LC-BiDFE is illustrated in Figure A.6. For an SER value of 10^{-3}, the iterated LC-BiDFE (with soft decisions) offers an additional performance improvement of nearly 0.3 dB over the soft LC-BiDFE with the "Suboptimal" tap settings. The label "LC-BiDFE : Suboptimal" refers to the use of the

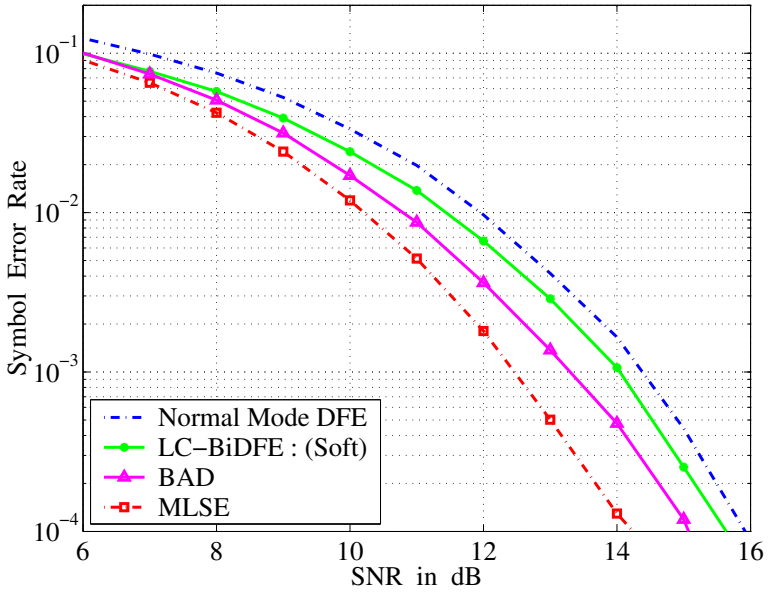

Figure A.5: SER performance comparison for a deep null channel with impulse response $\mathbf{c}_3 = [-0.3920 \;\; 0.0152 \;\; 0.8320 \;\; 0.0152 \;\; -0.3920]$

MMSE-DFE tap settings for the LC-BiDFE. As noted before, the iterative LC-BiDFE approach has a two-fold complexity when compared to the LC-BiDFE.

A.4 Fading Channel Simulation for MIMO LC-BiDFE

In this section, we consider the equalization of a quasi-static frequency-selective fading MIMO channel. The channel impulse response between each input and output of the system is assumed to be independent. Simulations were performed for an ensemble of 4×4 MIMO multipath channels, randomly generated using the urban environment power profile illustrated in Figure A.1. Each of the channel impulse responses were normalized to possess a unit norm. The source symbols were selected from a BPSK source constellation and perfect channel knowledge was assumed at the receiver.

Unbiased MMSE-DFE coefficients with $L_f = 16$ and $L_b = 10$ were used for the

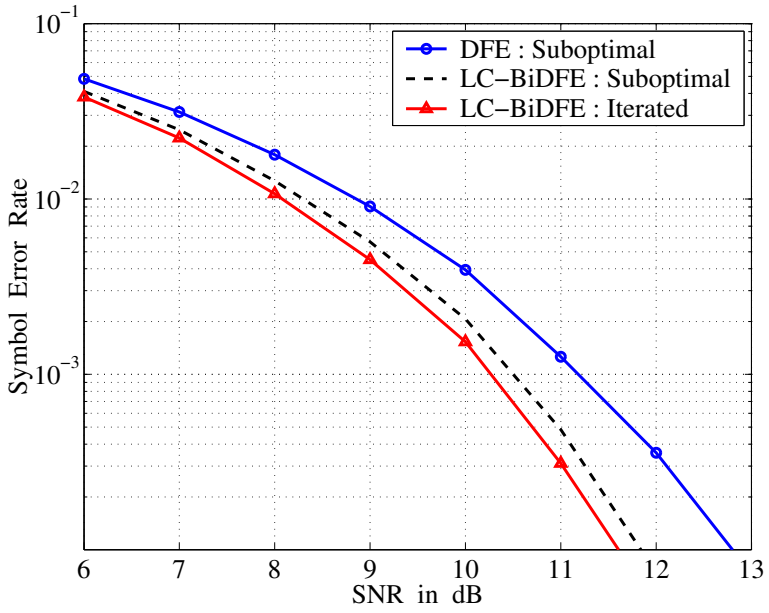

Figure A.6: SER performance of an iterated LC-BiDFE for the channel with impulse response $\mathbf{c}_4 = [-0.2053 \;\; 0.1300 \;\; 0.9391 \;\; 0.1300 \;\; -0.2053]$

MIMO DFEs for both the normal mode and time-reversal mode operation. The order of detection for normal mode DFE was, $\mathbf{ord}_1 = [s_1(n),\ s_2(n),\ s_3(n),\ s_4(n)]$, while for the time-reversal mode DFE, $\mathbf{ord}_2 = [s_4(n),\ s_3(n),\ s_2(n),\ s_1(n)]$. The SER performance curves (averaged over all the channel realizations) for a normal mode DFE, time-reversal mode DFE and the LC-BiDFE are compared in Figure A.7 for the four users $s_1(n)$, $s_2(n)$, $s_3(n)$ and $s_4(n)$, respectively. These simulation results do not invoke the assumption 2.5, and hence incorporate the effect of error propagation, i.e., decision feedback was assumed. The MIMO LC-BiDFE (with soft decisions) offers a performance improvement of about 1.5 dB, at an SER of 10^{-3}, over the time-reversal mode MIMO DFE for all the users. These results demonstrate the viability of the LC-BiDFE in providing performance improvements for the MIMO channel equalization problem.

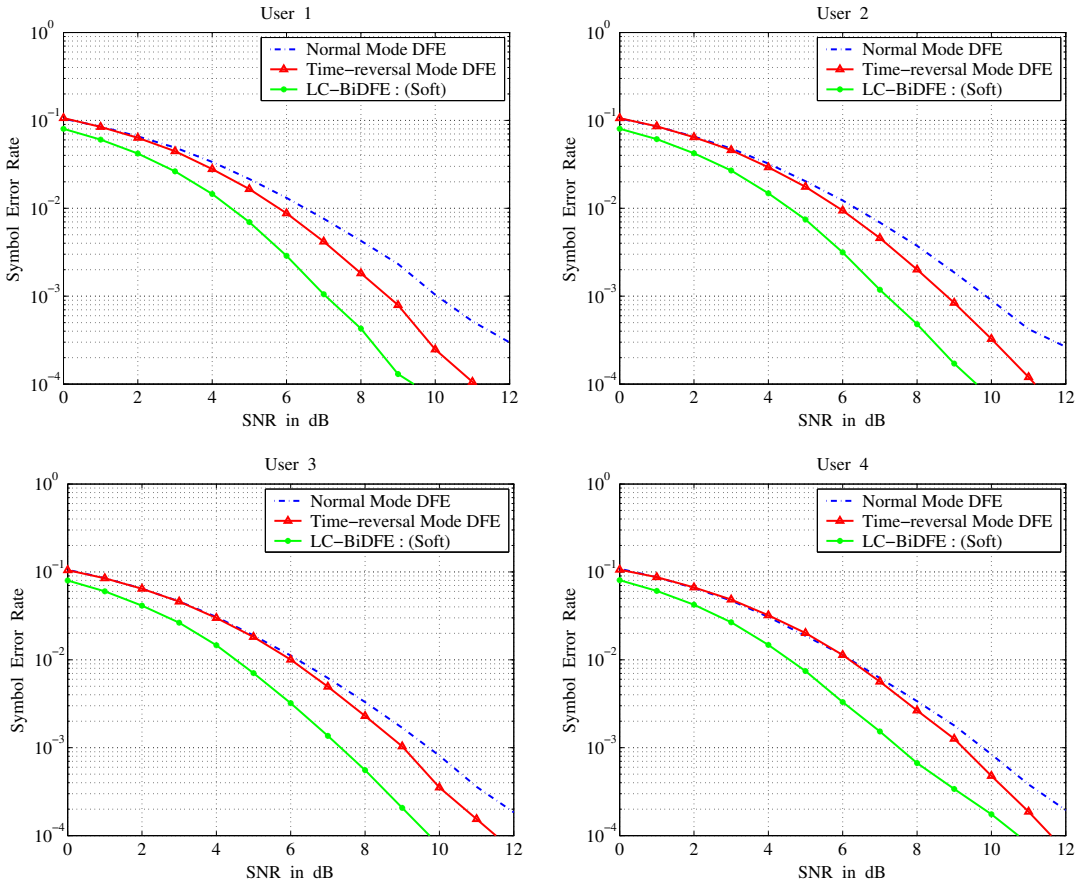

Figure A.7: SER performance comparison of users 1-4 for a MIMO fading channel environment with $M = 4,\ P = 4$

Appendix B

Training Sequences for EDGE

In chapter 6, the criterion of optimality for training sequences and a few search techniques for optimal and near-optimal training sequences were discussed. In [89], the authors propose the employment of multiple antennas to an EDGE system to increase system throughput. To enable channel estimation in such a MIMO system, near-optimal training sequences suitable for an EDGE system are determined and are tabulated in this appendix.

First, we consider an EDGE system with 2 transmit antennas. Based on the various design metrics described in Section 6.3, a training length of N = 26 was chosen for the 2 transmit antenna case. From equation (6.17), for this choice of training sequence length, the maximum identifiable number of the channel impulse response taps is 9. A typical urban channel impulse response, with pedestrian mobility (about 3 Km/h), has most of the energy concentrated in either 4 or 5 taps, and we make a conservative choice of $L = 7$ for the channel impulse response to be estimated. It should be noted that L, when used in the design of training sequences, is merely a design parameter and as long as the actual length of the channel impulse response to be estimated is less than the design parameter L, the designed training sequences will be near-optimal. The training

Table B.1: Training sequences for an EDGE system with $M = 2$

Antenna 1	**Antenna 2**	**Penalty Over Ideal Training**
0FB5D8F	293BE29	0.1599 dB
0391483	251F725	0.1599 dB
3785377	0BB9F4B	0.1599 dB
3BB287B	0B4188B	0.1599 dB
1D2F9DD	21135E1	0.1599 dB
11182D1	21EB221	0.1599 dB
2F0A6EF	1773E97	0.1599 dB
3DD943D	05A0C45	0.1599 dB

symbols were determined based on the random search method, described in [13] and in subsection 6.4.2, and were restricted to be from a BPSK constellation. A few pairs of these near-optimal training sequences are shown in hexadecimal format in Table B.1. The most-significant-bit (MSB) of the hexadecimal representation corresponds to the first symbol of the training sequence. The bit 1 corresponds to the symbol "+1" and the bit 0 to the symbol "-1". The penalty incurred, in terms of the loss in effective SNR due to channel estimation, by the sub-optimal training sequences when compared to the ideal training sequences is also tabulated in Table B.1 and is as small as 0.16 dB. This penalty is computed as the additional loss in effective SNR caused by the choice of the sub-optimal training sequences over the ideal training sequences and is given by (recall equations (6.10) and (6.19)),

$$\text{Penalty over ideal training} = 10 \log \left[\frac{1 + \text{tr}\left\{ \left(\mathbf{S}^H \mathbf{S} \right)^{-1} \right\}}{1 + \frac{ML}{(N-L+1)}} \right]. \qquad \text{(B.1)}$$

Figure B.1 illustrates[1] the loss incurred by the sub-optimal training sequences, when estimating channel impulse responses of varying lengths, over the ideal training sequences. The penalty over ideal training becomes significant (about 0.8 dB) only when the channel impulse response has a length of 8, which is beyond the design choice of $L = 7$.

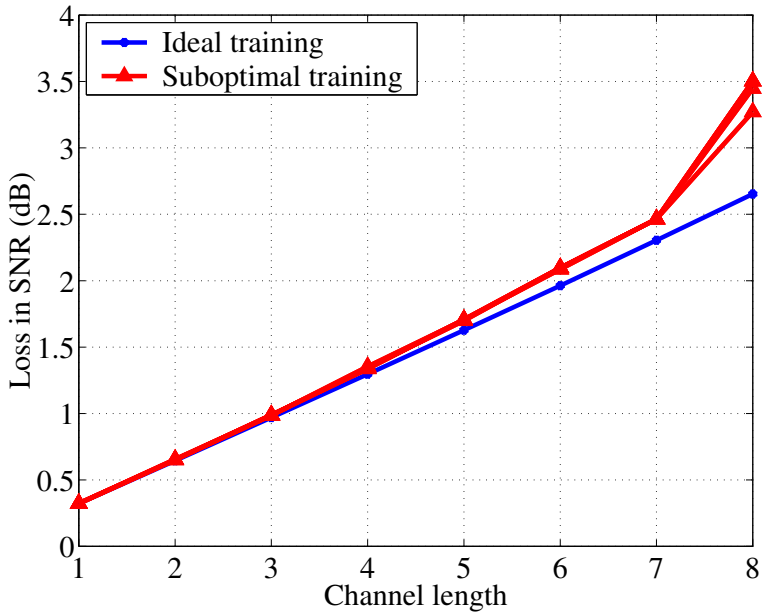

Figure B.1: Loss due to channel estimation, $M = 2$

Now we consider the use of four antennas at the transmitter, namely $M = 4$. For this choice of a transmit antennas it was shown in [89] that a training sequence with length $N = 36$ is suitable in the sense that it nearly maximizes the average throughput of the system while affording identifiability of the channel impulse response. For this choice of a training sequence length, from equation (6.17) we note that a channel impulse response with a maximum of 7 taps can be identified. For the training sequence design, a delay spread of $L = 5$ was assumed. This choice was motivated by the fact that, for a

[1] See [19] for Matlab source code

Table B.2: Training sequences for an EDGE system with $M = 4$

Antenna 1	**Antenna 2**	**Antenna 3**	**Antenna 4**	**Penalty Over Ideal Training**
0A7076510	7076510A7	76510A707	510A70765	0.0738 dB
2F9291822	9291822F9	91822F929	822F92918	0.0738 dB
517A46305	7A4630517	4630517A4	30517A463	0.0738 dB
C2D45980C	D45980C2D	5980C2D45	80C2D4598	0.1433 dB
2D8B8E402	8B8E402D8	8E402D8B8	402D8B8E4	0.1349 dB
B6E05238B	E05238B6E	5238B6E05	38B6E0523	0.1166 dB
59B80A8E5	B80A8E59B	0A8E59B80	8E59B80A8	0.1191 dB
CC876AEBC	876AEBCC8	6AEBCC876	EBCC876AE	0.1110 dB

TU3 (typical urban environment with a mobility of 3 Km/h) channel impulse response most of the energy is concentrated in either 4 or 5 taps. As in the 2 transmit antenna case, the training symbols were restricted to be from a BPSK constellation. A few sets of sub-optimal training sequences are tabulated in hexadecimal format in Table B.2. Further, the penalty incurred due to the choice of these sequences over the ideal training sequences is also tabulated in Table B.2. These training sequences were obtained using the cyclic shift search method described in subsection 6.4. For the design choice of $L = 5$, the loss incurred in effective SNR due to channel estimation by the sub-optimal training sequences over the ideal training sequences is at most 0.14 dB. Figure B.2 illustrates[2] the loss incurred by the sub-optimal training sequences, when estimating channel impulse responses of varying lengths, over the ideal training sequences. The

[2]See [19] for Matlab source code

incurred penalty is in the order of 1 dB, only when the channel impulse response has 7 taps, which is higher than the design choice of $L = 5$.

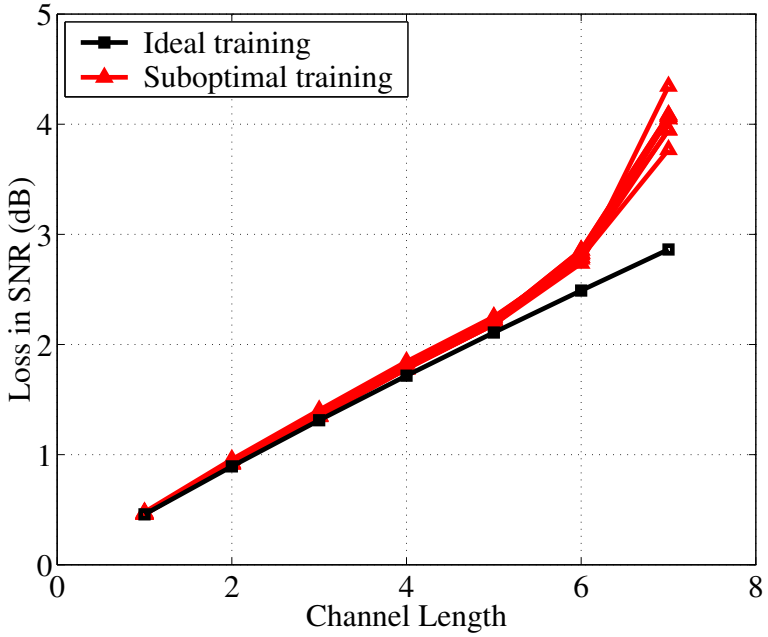

Figure B.2: Loss due to channel estimation, $M = 4$

Appendix C

Birding the Cayuga Lake Basin

The Cayuga basin is loosely demarcated by the borders of the watershed that flow into the Cayuga lake. The Cayuga basin bio-region plays host to a number of bird species - breeders, migrants and winter irruptives. Numerous birding hot spots in the Cayuga basin provide ample opportunity to a birder to see and appreciate the over 300 species of birds that have been sighted at one time or other in the past. A listing of the birding hot spots in the Cayuga basin can be found in [66]. The Cayuga bird club website [23] is a useful resource to learn about local birds and birding localities. The 244 bird species sighted by the author in the Cayuga lake basin, between July 1999 and May 2002, are listed in this appendix.

Table C.1: List of bird species seen in the Cayuga lake basin

Common loon	Red-throated loon	Pied-billed grebe
Horned grebe	Eared grebe	Double-crested cormorant
Least bittern	Great blue heron	Great egret
Green heron	Black-crown. night heron	Wood stork

Table C.1 (continued)

Turkey Vulture	Snow goose	Canada goose
Brant	Mute Swan	Tundra swan
Wood duck	Gadwall	American wigeon
American black duck	Mallard	Blue-winged teal
Northern shoveler	Northern pintail	Green-winged teal
Canvasback	Redhead	Ring-necked duck
Greater scaup	Lesser scaup	Surf scoter
Black scoter	Long-tailed duck	Bufflehead
Common goldeneye	Hooded merganser	Common merganser
Red-breasted merganser	Ruddy duck	Osprey
Bald eagle	Northern harrier	Sharp-shinned hawk
Cooper's hawk	Broad-winged hawk	Red-tailed hawk
Rough-legged hawk	Golden eagle	American kestrel
Merlin	Peregrine falcon	Ring-necked pheasant
Ruffed grouse	Wild turkey	Virginia rail
Sora	Common moorhen	American coot
Sandhill crane	Black-bellied plover	American golden plover
Semipalmated plover	Piping plover	Killdeer
Greater yellowlegs	Lesser yellowlegs	Solitary sandpiper
Spotted sandpiper	Upland sandpiper	Hudsonian godwit
Marbled godwit	Ruddy turnstone	Semipalmated sandpiper
Least sandpiper	Pectoral sandpiper	Purple sandpiper
Dunlin	Stilt sandpiper	Buff-breaster sandpiper

Table C.1 (continued)

Short-billed dowitcher	Long-billed dowitcher	Common snipe
American woodcock	Wilson's phalarope	Red-necked phalarope
Bonaparte's gull	Ring-billed gull	Greater Black-backed gull
Iceland gull	Herring gull	Lesser black-backed gull
Caspian tern	Common tern	Forster's tern
Black tern	Long-billed murrelet	Rock dove
Mourning dove	Yellow-billed cuckoo	Eastern screech owl
Great-horned owl	Snowy owl	Long-eared owl
Short-eared owl	Northern Saw-whet owl	Common nighthawk
Whip-poor-will	Chimney swift	Ruby-thr. hummingbird
Belted kingfisher	Red-headed woodpecker	Red-bellied woodpecker
Yellow-bellied sapsucker	Downy woodpecker	Hairy woodpecker
Northern flicker	Pileated woodpecker	Olive-sided flycatcher
Eastern wood-peewee	Yellow-bellied flycatcher	Acadian flycatcher
Alder flycatcher	Willow flycatcher	Least flycatcher
Eastern phoebe	Great-crested flycatcher	Western kingbird
Eastern kingbird	Northern shrike	Yellow-throated vireo
Blue-headed vireo	Warbling vireo	Philadelphia vireo
Red-eyed vireo	Blue jay	American crow
Fish Crow	Common raven	Horned lark
Purple martin	Tree swallow	N. rough-winged swallow
Bank swallow	Cliff swallow	Barn swallow
Black-capped chickadee	Tufted titmouse	Red-breasted nuthatch

Table C.1 (continued)

White-breasted nuthatch	Brown creeper	Carolina wren
House wren	Winter wren	Marsh wren
Golden-crowned kinglet	Ruby-crowned kinglet	Blue-gray gnatcatcher
Eastern bluebird	Veery	Gray-cheeked thrush
Swainson's thrush	Hermit thrush	Wood thrush
American robin	Gray catbird	Northern mockingbird
Brown thrasher	European starling	American pipit
Bohemian waxwing	Cedar waxwing	Blue-winged warbler
Golden-winged warbler	Tennessee warbler	Orange-crowned warbler
Nashville warbler	Northern parula	Yellow warbler
Chestnut-sided warbler	Magnolia warbler	Cape-May warbler
Black-thr. blue warbler	Yellow-rumped warbler	Black-thr. green warbler
Blackburnian warbler	Pine warbler	Prairie warbler
Palm warbler	Bay-breasted warbler	Blackpoll warbler
Cerulean warbler	Black-and-white warbler	American redstart
Worm-eating warbler	Ovenbird	Northern waterthrush
Louisiana waterthrush	Mourning warbler	Common yellowthroat
Hooded warbler	Wilson's warbler	Canada warbler
Yellow-breasted chat	Scarlet tanager	Eastern Towhee
American tree sparrow	Chipping sparrow	Field sparrow
Vesper sparrow	Savannah sparrow	Grasshopper sparrow
Henslow's sparrow	Fox sparrow	Song sparrow
Lincoln's sparrow	Swamp sparrow	White-throated sparrow

Table C.1 (continued)

White-crowned sparrow	Dark-eyed junco	Lapland longspur
Snow bunting	Northern cardinal	Rose-breasted grosbeak
Indigo bunting	Bobolink	Red-winged blackbird
Eastern meadowlark	Western meadowlark	Rusty blackbird
Brown-headed cowbird	Common grackle	Orchard oriole
Baltimore oriole	Pine grosbeak	Purple finch
House finch	White-winged crossbill	Common redpoll
Pine siskin	American goldfinch	Evening grosbeak
House sparrow		

Bibliography

[1] N. Al-Dhahir and J. M. Cioffi. MMSE Decision-Feedback Equalizers: Finite-Length Results. *IEEE Trans. Information Theory*, 41(4):961–975, July 1995.

[2] N. Al-Dhahir and A. H. Sayed. The Finite-Length Multi-Input Multi-Output MMSE-DFE. *IEEE Trans. Signal Processing*, 48(10):2921–2936, October 2000.

[3] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Communications*, 16(8):1451–1458, October 1998.

[4] S. A. Altekar and N. C. Beaulieu. Upper bounds to the error probability of decision feedback equalization. *IEEE Trans. Information Theory*, 39(1):145–156, January 1993.

[5] S. Ariyavisitakul. A decision-feedback equalizer with selective time-reversal operation for high-rate indoor radio communication. In *Proc. GLOBECOM '90*, pages 2035–2039, December 1990.

[6] S. Ariyavisitakul. A Decision Feedback Equalizer with Time-Reversal Structure. *IEEE J. Sel. Areas Communications*, 10(3):599–613, April 1992.

[7] S. Ariyavisitakul and G. M. Durant. A broadband wireless packet technique based on coding, diversity and equalization. *IEEE Communications Magazine*, 36(7):110–115, July 1998.

[8] S. Ariyavisitakul and Y. Li. Joint coding and decision feedback equalization for broadband wireless channels. In *Proc. IEEE Vehicular Technology Conference*, pages 2256–2261, May 1998.

[9] M. Austin. Decision feedback equalization for digital communication over dispersive channels. *MIT Research Laboratory of Electronics Technical Report 437*, August 1967.

[10] J. Balakrishnan. Mitigation of Error Propagation in Decision Feedback Equalization. Master's thesis, Cornell University, Aug 1999.

[11] J. Balakrishnan and C. R. Johnson, Jr. Time-Reversal Diversity in Decision Feedback Equalization. In *Proc. of Allerton Conf. on Comm. Control and Computing*, Monticello, IL, USA, October 2000.

[12] J. Balakrishnan and C. R. Johnson, Jr. Time-Reversal Decision Feedback Equalizer: Infinite Length Results. In *Proc. of Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, November 2001.

[13] J. Balakrishnan, M. Rupp, and H. Viswanathan. Optimal Channel Training for Multiple Antenna Systems. In *Proc. of Conf. on Multi-access, Mobility and Teletraffic*, Duck Key, FL, USA, December 2000.

[14] J. Balakrishnan, H. Viswanathan, and C. R. Johnson, Jr. Decision Device Optimization for Soft Decision Feedback Equalization. In *Proc. of CISS 2000*, Princeton, NJ, USA, March 2000.

[15] G. Barriac and U. Madhow. PASIC: A New Paradigm for Low-Complexity Multiuser Detection. In *Proc. of CISS 2001*, Baltimore, MD, USA, March 2001.

[16] N. C. Beaulieu. Bounds on Recovery Times of Decision Feedback Equalizers. *IEEE Trans. Communications*, 42(10):2786–2794, October 1994.

[17] C. A. Belfiore and J. H. Park, Jr. Decision feedback equalization. *Proc. IEEE*, 67(8):1143–1156, August 1979.

[18] L. Bomer and M. Antweiler. Perfect N-phase sequences and arrays. *IEEE J. Sel. Areas Communications*, 10:782–789, May 1992.

[19] Broadband Adaptive Receiver Design Project, Cornell University, http://bard.ece.cornell.edu/matlab/balakrishnan/index.html. *MATLAB Source Code*.

[20] A. M. Chan and G. W. Wornell. A Class of Block-Iterative Equalizers for Intersymbol Interference Channels. In *Proc. Int. Conf. on Communications*, New Orleans, LA, June 2000.

[21] A. M. Chan and G. W. Wornell. A Class of Block-Iterative Equalizers for Intersymbol Interference Channels: Fixed Channel Results. *IEEE Trans. Communications*, 49(11):1966–1976, November 2001.

[22] A. Chevreuil and L. Vandendorpe. MIMO MMSE-DFE: A General Framework. In *Proc. IEEE Workshop on Statistical Signal and Array Processing*, pages 368–371, Portland, OR, USA, September 1998.

[23] Cornell Lab of Ornithology, http://www.birds.cornell.edu/cayugabirdclub/. *The Cayuga Bird Club*.

[24] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud. Least sum of squared errors (LSSE) channel estimation. *IEE Proc. F*, 138(4):371–378, 1991.

[25] J. E. Mazo D. L. Duttweiler and D. G. Messerschmitt. An upper bound on the error probability in decision feedback equalization. *IEEE Trans. Information Theory*, 20(4):490–497, July 1974.

[26] E. de Carvalho and D. T. M. Slock. Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation. In *Proc. Signal Processing Advances in Wirelss Comm.*, Paris, France, 1997.

[27] E. de Carvalho and D. T. M. Slock. Burst Mode Non-Causal Decision-Feedback Equalizer based on Soft Decisions. In *Proc. VTC '98*, pages 414–418, Ottawa, Canada, May 1998.

[28] A. M. de Oliveira Duarte and J. J. O'Reilly. Simplified technique for bounding error statistics for DFB receivers. *Proc. IEEE*, 132(7):567–575, December 1985.

[29] K. Dogancay and R. A. Kennedy. Blind detection of equalization errors in communication systems. *IEEE Trans. Information Theory*, 43(2):469–482, March 1997.

[30] A. Duel-Hallen. Equalizers for Multiple Input/Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences. *IEEE J. Sel. Areas Communications*, 10(3):630–639, April 1992.

[31] A. Duel-Hallen and C. Heegard. Delayed Decision Feedback Sequence Estimation. *IEEE Trans. Communications*, 37(5):428–435, May 1989.

[32] R. Steele (Ed). *Mobile Radio Communications*. IEEE Press, 1992.

[33] ETSI. Digital Cellular Telecommunications System. *GSM Specification 05.02 Version 8.3.0*, ETR 364, 1999.

[34] W. Van Etten. Maximum Likelihood Receiver for Multiple Channel Transmission Systems. *IEEE Trans. Communications*, COM-24(2):276–283, February 1976.

[35] M. V. Eyuboğlu and S. U. Qureshi. Reduced-state sequence estimation with set partitioning and decision feedback. *IEEE Trans. Communications*, 36:13–20, January 1988.

[36] A. Fertner. Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation. *IEEE Trans. Signal Processing*, 46(7):1872–1877, July 1998.

[37] G. J. Foschini. Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multi-Element Antennas. *Bell Labs Technical Journal*, 1(1):41–59, Autumn 1996.

[38] G. J. Foschini and M. J. Gans. On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas. *Wireless Personal Communications*, pages 311–335, March 1998.

[39] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolaniansky. Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multi-element Arrays. *IEEE J. Sel. Areas Communications*, 17(11):1841–1852, November 1999.

[40] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson. EDGE: Enhanced data rates for GSM and TDMA/136 evolution. *IEEE Personal Communications Mag.*, June 1999.

[41] G. D. Forney, Jr. Maximum-likelihood sequence estimation of digital sequences in the presence of inter symbol interference. *IEEE Trans. Information Theory*, 18(2):989–995, May 1972.

[42] G. D. Forney, Jr. The Viterbi Algorithm. *Proc. IEEE*, 61(3):268–278, March 1973.

[43] A. Gersho and T. L. Lim. Adaptive Cancellation of Intersymbol Interference for Data Transmission. *Bell Systems Technical Journal*, 60(11):1997–2021, November 1981.

[44] M. Ghosh. Analysis of the MMSE-DFE with Error Propagation. In *IEEE GLOBECOM Communication Theory Mini-Conf*, pages 85–89, November 1997.

[45] M. Ghosh. Blind decision feedback equalization for terrestrial television receivers. *Proc. IEEE*, 86(10):2070–2081, October 1998.

[46] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. Communications*, 29(2):774–780, August 1972.

[47] S. Haykin. *Communication Systems*. Wiley, Newyork, third edition, 1994.

[48] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Upper Saddle River, NJ, 1996.

[49] G. K. Kaleh. Channel Equalization for Block Transmission Systems. *IEEE J. Sel. Areas Communications*, 13(1):110–121, January 1995.

[50] Y. Kamio and S. Sampei. Performance of reduced complexity DFE using bidirectional equalizing in land mobile communications. In *Proc. of VTC '92*, pages 372–375, Denver, CO, USA, May 1992.

[51] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, 1993.

[52] R. A. Kennedy and B. D. O. Anderson. Recovery times of decision feedback equalizers on noiseless channels. *IEEE Trans. Communications*, 35(10):1012–1021, October 1987.

[53] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead. Tight bounds on the error probabilities of decision feedback equalizers. *IEEE Trans. Communications*, 35(10):1022–1028, October 1987.

[54] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead. Channels leading to rapid error recovery for decision feedback equalizers. *IEEE Trans. Communications*, 37(11):1126–1135, November 1989.

[55] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead. Stochastic analysis of non-adaptive decision feedback equalizers. In *Proc. ISSPA97*, pages 842–847, Brisbane, Australia, August 1997.

[56] R. L. Kosut, W. Chung, C. R. Johnson, Jr., and S. P. Boyd. On Achieving Reduced Error Propagation Sensitivity in DFE Design via Convex Optimization. In *Proc. IEEE Conf. on Decision and Control*, Sydney, Australia, December 2000.

[57] W. U. Lee and F. S. Hill, Jr. A maximum-likelihood sequence estimator with decision-feedback equalization. *IEEE Trans. Communications*, 25(9):971–979, September 1977.

[58] S. Li and M. P. Fitz. On Reduced State Equalizers. In *Proc. of Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, November 2001.

[59] E. Lindskog and A. Paulraj. A Transmit Diversity Scheme for Channels with Intersymbol Interference. In *Proc. of ICC 2000*, pages 307–311, New Orleans, LA, USA, June 2000.

[60] Y. Liu, M. Wallace, and J. W. Ketchum. A Soft-Output Bidirectional Decision Feedback Equalization Technique for TDMA Cellular Radio. *IEEE J. Sel. Areas Communications*, 11(7):1034–1045, September 1993.

[61] R. W. Lucky. Automatic Equalization for Digital Communications. *Bell Systems Technical Journal*, 44:547–588, April 1965.

[62] M. Marandian, M. Salehi, J. G. Proakis, and F. Blackmon. Iterative Decision Feedback Equalizer for Frequency Selective Fading Channels. In *2001 Conf. on Information Sciences and Systems*, Baltimore, MD, USA, March 2001.

[63] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities*. Dover Publications Inc., Newyork, 1964.

[64] C. S. McGahey, A. C. Singer, and U. Madhow. BAD: A Bi-directional Arbitrated Decision Feedback Equalizer. In *Proc. of CISS 2000*, Princeton, NJ, USA, March 2000.

[65] N. C. McGinty and R. A. Kennedy. Reduced-State Sequence Estimator with Reverse-Time Structure. *IEEE Trans. Communications*, 45(3):265–268, March 1997.

[66] D. W. McIlroy and C. R. Smith. *Birding in the Cayuga Lake basin*. Cornell Lab of Ornithology, Ithaca, NY, 1993.

[67] D. Messerschmitt. A Geometric theory of intersymbol interference. Part I: Zero-forcing and decision feedback equalization. *Bell System Technical Journal*, 52(10):1483–1519, November 1973.

[68] A. Milewski. Periodic sequences with optimal properties for channel estimation and fast start-up equalization. *IBM J. of Research and Development*, 27:426–431, 1983.

[69] P. Monsen. *Linear equalization for digital transmission over noisy dispersive channels*. PhD thesis, Columbia University, Jun 1970.

[70] P. Monsen. Feedback equalization for fading dispersive channels. *IEEE Trans. Information Theory*, 17:56–64, January 1971.

[71] R. V. Nee and R. Prasad. *OFDM for Wireless Multimedia Communications*. Artech House Publishers, 2000.

[72] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, first edition, 1989.

[73] J. J. O'Reilly and A. M. de Oliveira Duarte. Error propagation in decision feedback equalizers. *Proc. IEEE*, 132(7):561–566, December 1985.

[74] R. Price. Nonlinearly feedback-equalized PAM vs. capacity for noisy filter channels. In *Rec. Int. Conf. Communications*, Philadelphia, PA, 1972.

[75] J. G. Proakis. Adaptive Non Linear Filtering Techniques for Data Transmission. *IEEE Symposium on Adaptive Processes*, pages XV.2.1–5, 1970.

[76] J. G. Proakis. *Digital Communications*. McGraw-Hill, New York, NY, third edition, 1995.

[77] D. Raphaeli. Decoding Algorithms for Noncoherent Trellis Coded Modulation. *IEEE Trans. Communications*, 44(3):312–323, March 1996.

[78] D. Raphaeli and T. Kaitz. A Reduced-Complexity Algorithm for Combined Equalization and Decoding. *IEEE Trans. Communications*, 48(11):1797–1807, November 2000.

[79] M. Rupp. FAST Implementation of the LMS Algorithm. In *Proc. EUSIPCO 2000*, Tampere, September 2000.

[80] J. Salz. Optimum mean-square decision feedback equalization. *Bell System Technical Journal*, 52(8):1341–1373, October 1973.

[81] D. T. M. Slock and E. de Carvalho. Burst Mode Non-Causal Decision-Feedback Equalization and Blind MLSE. In *Proc. GLOBECOM 96*, London, Great Britain, November 1996.

[82] D. T. M. Slock and E. de Carvalho. Matched Filter Bounds for Reduced-Order Multichannel Models. In *Proc. GLOBECOM 96*, pages 77–81, London, United Kingdom, November 1996.

[83] H. Suzuki. Performance of a new adaptive diversity-equalization for digital mobile radio. *Electron. Letters*, 26(10):626–627, May 1990.

[84] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communications. *IEEE Trans. Information Theory*, 44(2):744–765, March 1998.

[85] C. Tellambura, Y. J Guo, and S. K. Barton. Channel estimation using aperiodic training sequences. *IEEE Communication Letters*, 2(5):140–142, 1998.

[86] C. Tellambura, M. G. Parker, Y. J Guo, S. J. Sheperd, and S. K. Barton. Optimal sequences for channel estimation using discrete Fourier transform techniques. *IEEE Trans. Communications*, 47(2):230–237, 1999.

[87] M. Tomlinson. New Automatic Equalizers Employing Modulo Arithmetic. *Electron. Letters*, 7(5-6):138–139, March 1971.

[88] G. L. Turin. An introduction to digital matched filters. *Proc. IEEE*, 64:1092–1112, July 1976.

[89] H. Viswanathan and J. Balakrishnan. Space-time Signaling for High Data Rates in EDGE. *To appear in IEEE Trans. on Vehicular Technology*.

[90] J. H. Winters, J. Salz, and R. D. Gitlin. The impact of antenna diversity on the capacity of wireless communication systems. *IEEE Trans. Communications*, 47(7):1073–1083, July 1999.