MITIGATION OF ERROR PROPAGATION IN DECISION FEEDBACK EQUALIZATION

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Jaiganesh Balakrishnan

August 1999

ABSTRACT

A major obstacle to reliable digital communication is inter-symbol interference (ISI) and noise, encountered in transmission over dispersive channels. Thus, equalization of the channel at the receiver is necessary. There has been immense interest in the research community, since the early 1970s, on decision feedback equalizers (DFE) which are very effective and relatively easy to implement. Unfortunately, DFE performance is hampered by error propagation. Though this phenomenon has been known for a long time, only a modest amount of work has been done on characterizing it and very little work has been done on mitigating error propagation.

The bulk of DFE research has been devoted to proposal and performance study of various modifications of the basic DFE. In this tradition, a soft decision feedback equalization scheme has been proposed in this thesis, where the hard-limiter of a conventional DFE is replaced with a soft decision device. The intention of this modification is mitigation of error propagation. The mean squared error (MSE) optimized soft decision device is derived here using various simplifying assumptions. Simulation results demonstrating performance improvements are provided. Further, a bit error rate (BER) optimized soft decision device is derived for a BPSK modulation scheme. The joint optimization of the soft DFE decision device and the DFE taps is discussed for the special case of a source with Gaussian distribution.

Performance evaluation of the soft DFE based on a multi-level quantization approximation is discussed. The error dynamics of the system are modeled using a

Markov chain and the corresponding state occupancy probabilities are computed. MSE and BER of the soft DFE are calculated from these probabilities. A few sub-optimal implementations, namely the multi-level quantization scheme and the piece-wise linear approximation, of the soft DFE decision device are considered. Simulation results on performance comparison of soft DFE with the conventional DFE, for measured microwave and broadband wireless channels, are provided.

In practical communication systems, error correction coding is a critical component needed to achieve reduced bit error rates. Most coding schemes are designed to correct random errors and perform poorly in the presence of ISI. This has motivated research on simple but effective joint equalization and decoding strategies. Combining DFE with error correction coding is a challenging problem due to the presence of error propagation. The bursty nature of DFE errors causes performance loss in the decoder.

A turbo DFE strategy, based on the turbo-equalization principle, has been proposed in this thesis. In this scheme equalization and decoding are performed iteratively. The more reliable decisions at the output of the decoder are fed back as inputs to the feedback filter of the DFE, thereby reducing the probability of occurrence of primary errors and reducing the likelihood of error propagation. Simulation results are provided for a few examples.

Biographical Sketch

Jaiganesh Balakrishnan was born in Madras, the largest city in South India in the year 1976. He spent the first 17 years of his life in Chromepet, a suburb of Madras. He did his schooling at N.S.N Matriculation Higher Secondary School, Chromepet. He joined the Bachelor of Technology program in Electrical Engineering at the Indian Institute of Technology, Madras in 1993. He entered graduate school at Cornell University in the fall of 1997. In his spare time, Jaiganesh enjoys birding, reading, watching Tamil movies and playing cricket.

To Tamil

யாமறிந்த மொழிகளிலே தமிழ்மொழி போல் இனிதாவது எங்கும் காணோம்,
பாமரராய், விலங்குகளாய், உலகனைத்தும் இகழ்ச்சிசொலப் பான்மை கெட்டு
நாமமது தமிழரெனக்கொண்டு இங்கு வாழ்ந்திடுதல் நன்றோ? சொல்லீர் !
தேமதுரத் தமிழோசை உலகமெல்லாம் பரவும்வகை செய்தல் வேண்டும்.

சுப்ரமணிய பாரதியார் (1882 - 1921)

Acknowledgements

I would like to thank my advisor Dr. C.R. Johnson, Jr., for his guidance and encouragement. His comments and suggestions were very useful for the improvement of my technical writing style. I would like to thank Dr. Lang Tong for being on my committee and for his enthusiastic teaching. I would also like to thank him for his various suggestions on my thesis. I thank all my teachers at Cornell and I.I.T, Madras for their efforts.

I am grateful to 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 take this opportunity to thank Rajeev, my supervisor at Lucent Technologies, for his guidance and help. I thank Harish, my mentor and friend at Lucent, without whom this thesis would not have been possible. Thanks also for assistance from NSF Grant ECS-9528363.

I thank Dr. S. Sundaram for his encouragement and advice throughout my life and Dilip for being my friend, philosopher and guide. Thanks goes to all my house mates, research group members and friends for their support and encouragement. I thank Ashok for constantly enlightening me about various aspects of life.

I utilize this opportunity to express my deep gratitude to my parents and sister who have sacrificed so much for my sake.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Future generation wireless communication technology promises an increase in data rate capabilities of present radio links. As we go for higher rates the frequency distortion caused by the channel becomes worse. The distortion caused by the resulting overlap of received signals is called Inter Symbol 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.

A powerful technique for combating ISI is the Maximum-Likelihood Sequence Detection (MLSD) [5], which minimizes probability of symbol error. This joint equalization and detection technique can be implemented with the trellis-based Viterbi algorithm [7]. 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 M and the number of interfering symbols contributing to ISI is L, the Viterbi algorithm computes M^{L+1} metrics for each new received symbol. When the channel has long memory (or large L) the MLSD approach becomes impractical. A number of schemes, e.g., [11, 18], generally known as reduced-state sequence estimation (RSSE), were proposed in the 1970s and 1980s in an attempt to approach the performance of MLSD at reduced complexity.

One suboptimal scheme that is widely used in practice, due to its simple implementation, is the linear equalizer [1]. This approach employs a linear transversal filter. This structure has a computational complexity that is a linear function of the channel dispersion length L. For a channel introducing only mild interference, the performance achievable by a conventional linear equalizer is often adequate. But, in suppressing the inter symbol interference, 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.

The basic limitation of a linear equalizer to cope with severe ISI has motivated a considerable amount of research into suboptimal nonlinear equalizers with low computational complexity such as the decision feedback equalizer. The classic structure of this receiver is shown in Figure 1.1. A linear feed-forward filter (FFF) processes the received signal and its output is sampled at time interval of T seconds. The sampler output is then the sum of a linear combination of several transmitted symbols plus a noise term. The receiver makes decisions every T seconds on a symbol-by-symbol basis. At the time a decision is to be made on the nth transmitted symbol the feedback filter (FBF) forms a weighted linear combination of the previous symbol decisions, assumed to be correct, and subtracts this combination from the FFF output. The result is then applied to a threshold device to determine the current symbol decision.

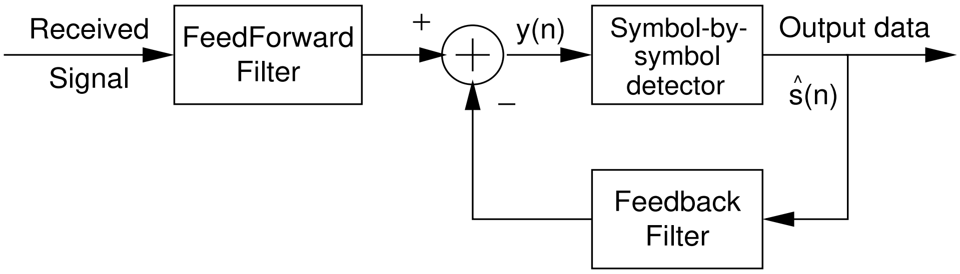

Figure 1.1: Structure of decision-feedback equalizer

The idea of using previous decisions to cope with the ISI problem was first introduced in 1967 by Austin [2] only two years after the development of the digital (adaptive) linear equalizer by Lucky [1]. The decision feedback receiver that minimizes the MSE between the input to the threshold device and the transmitted symbol was first obtained by Monsen [3]. The joint optimum MMSE transmitter and receiver was obtained by Salz [8]. Using linear space geometric arguments Messerschmitt [9] 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 [12], 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 [27]. Thus, developments in DFE design have continued over 30 years since its inception.

The DFE design has always been carried on 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, and the result is that errors tend to occur in bursts. The first residual-induced decision error, called a primary error, is fed back by the DFE 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 evaluation of the error propagation problem can be handled easily when a finite number L_b of past decisions enter into the current decision and L_b is very small. The system can then be modeled as a Markov chain with M^{L_b} states (where M is the number of distinct magnitude of errors), and when L_b is small the stationary probabilities can be determined numerically. For larger L_b it does not appear feasible to calculate the error probability exactly.

Some research, which began in the early 1970s shortly after invention of the DFE, has been done to simplify the analysis by clumping error states and providing upper bounds on the error probabilities. [10, 17, 24] use this approach to provide bounds on error propagation. [14, 15, 16, 25] 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 [20, 34].

However, little work has been done on mitigating error propagation in decision feedback equalizers. [32] proposes techniques for detecting decision errors in equalization schemes including DFEs. Several techniques, based on transmitter precoding [4], [6], were proposed in the early 1970s to remedy the error propagation problem in the DFE. However, these techniques require knowledge of the channel response by the transmitter, and thus might be inapplicable if the transmission is broadcast or if the channel is time-varying. [38] introduces 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. [36] proposes the use of a combination of soft decisions and delayed tentative decisions to cancel ISI for a joint convolutional coding and DFE scheme. The soft decision device is obtained by using a simplified maximum aposteriori probability (MAP) algorithm. This motivates our problem of finding simple and efficient schemes to be implemented at the receiver to mitigate error propagation in DFEs.

1.2 Thesis Outline

The thesis is organized as follows.

Chapter 2 briefly describes the system model and the notation used in the thesis. The problem of mitigating error propagation is formulated as an optimization of the decision device with a zero-order memory. Various approximations are made regarding residual inter-symbol interference and the noise statistics. The decision device input/output relation is optimized under the mean-squared error criterion and a general expression for the optimal decision device non-linearity is obtained. The decision box is derived for a few specific modulation schemes like QPSK, 8-PSK and 16QAM. A bit error rate optimization of the decision device for a BPSK modulation scheme is also derived. Simulation results comparing the performance of the soft DFE with the conventional DFE are provided. A special case of a source with Gaussian distribution is considered and the joint optimization of the soft DFE decision device and the DFE taps is discussed.

The difficulties involved in exact performance analysis of the soft DFE are discussed in Chapter 3. An approximate performance analysis based on a multi-level quantization of the soft DFE non-linearity is proposed. The error dynamics of the system are modeled using a Markov chain. The MSE and BER are calculated from the steady-state occupancy probabilities of the Markov chain. Numerical results demonstrating the accuracy of the approximations are shown.

The issue of implementing the decision device non-linearity is discussed in Chapter 4. Two suboptimal approximations, one based on multi-level quantization and the other on piece-wise linear approximation, for the decision box are proposed. Numerical results demonstrating the performance loss of the sub-optimal schemes are provided.

Simulation results illustrating performance comparisons of soft DFE with the conventional DFE are provided in Chapter 5, for a measured microwave channel and a few broadband wireless channel models specified by GSM.

Chapter 6 explores the idea of combining the DFE and the error-correcting decoder in an efficient manner. The turbo equalization principle, based on iteratively combining equalization and decoding, is reviewed. A turbo-like strategy is proposed for iteratively combining the DFE and the decoder. Simulation results comparing the performance of the turbo DFE with a conventional DFE and an ideal DFE are shown.

Finally, a summary of results and directions for future research are provided in Chapter 7

Chapter 2

Soft Decision Feedback Equalization

2.1 Motivation

The conventional DFE has always been designed under the assumption that the fed back decisions are correct. When error propagation is ignored, the hard-limiter is the optimal (under the minimum BER criterion) decision device. However, when the fed back decisions are incorrect, which is the case in reality, the hard-limiter is no longer the optimal decision device. This motivates the soft decision feedback approach, where the hard-limiter is replaced by a soft decision device. The soft decision device is derived in this chapter under a few simplifying assumptions. In particular: the channel has a finite impulse response(FIR) model, the optimum FFF and FBF are known, the channel noise and residual ISI have particular characteristics.

2.2 System Model

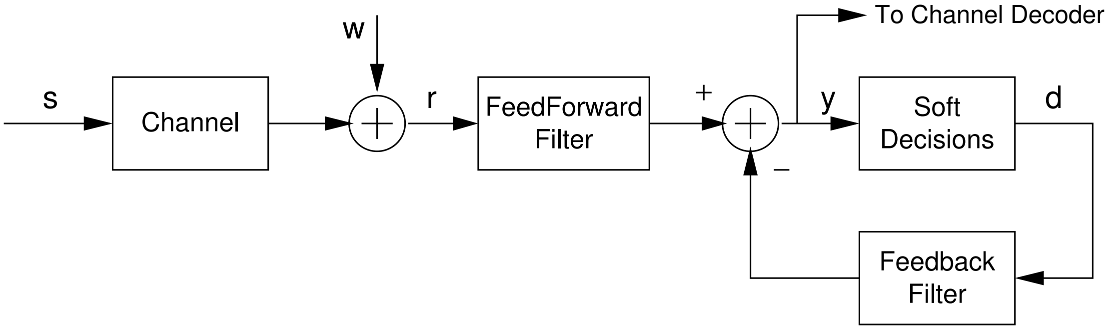

Figure 2.1: Digital baseband equivalent of system

Figure 2.1 depicts the digital baseband equivalent of our system. Typically, the information bits are mapped on to a signal constellation and transmitted. The transmitted symbol $s(n)$ belongs to a finite alphabet set $\mathcal{A}$. The channel incorporates the effects of frequency selective distortion caused by the physical medium and transmit pulse shaping. The transmitted signal is further corrupted by a complex additive white Gaussian noise (AWGN) at the receiver end. Now, we make a few simplifying assumptions.

- **Assumption 1:** The channel can be modeled as a linear FIR filter with complex coefficients. It is further assumed to be time invariant.

  Even if the channel varies with time, we assume that the equalizer can adapt to the channel variations.

The received signal $r(n)$ is given by,

$$r(n) = \begin{pmatrix} s(n) \\ s(n-1) \\ \vdots \\ s(n-L_c+1) \end{pmatrix}^T \mathbf{c} + w(n) \tag{2.1}$$

where $\mathbf{c} = [c(0)\ c(1)\ \ldots\ c(L_c-1)]^T$ is the channel impulse response with L_c taps and $w(n)$ is the additive white Gaussian noise. The received signal is equalized using a DFE which consists of a feed forward filter and a feedback filter. The feed forward filter has L_f taps and its impulse response is denoted by the vector $\mathbf{f} = [f(0)\ f(1)\ \ldots\ f(L_f-1)]^T$ and the feedback filter has L_b taps and its impulse response is given by the vector $\mathbf{b} = [b(1)\ b(2)\ \ldots\ b(L_b)]^T$. The combined channel-feed forward filter impulse response is given $\mathbf{h} = \mathbf{c} * \mathbf{f}$, where $*$ denotes the convolution operation. Further, let δ be the delay caused by the overall system, for some $\delta \geq 0$. $h(\delta)$ is referred to as the cursor. The input to the decision box is given by,

$$y(n) = \mathbf{s}^T\mathbf{h} + \mathbf{w}^T\mathbf{f} - \mathbf{d}^T\mathbf{b} \tag{2.2}$$

where $\mathbf{s} = [s(n+\delta)\ \ldots\ s(n+\delta-L_h+1)]^T$ is the transmitted symbol vector, $\mathbf{w} = [w(n)\ \ldots\ w(n-L_f+1)]^T$ is the noise vector, $\mathbf{d} = [d(n-1)\ \ldots\ d(n-L_b)]^T$ is the vector containing the past soft decisions. The feedback filter subtracts the contribution of the L_b taps of $\mathbf{h}$ that follow the cursor.

- **Assumption 2:** The feed-forward and feedback filter taps are known at the receiver. Further, let us assume that $\mathbf{b} = [h(\delta+1)\ h(\delta+2)\ \ldots\ h(\delta+L_b)]^T$ and $h(\delta) = 1$.

  In most applications the FFF and FBF taps are optimized under either the

MMSE or the zero-forcing criterion (ignoring the effects of error propagation). Since our goal is to optimize only the decision device, we make the above assumption. Various techniques for estimating the optimal DFE taps have been extensively discussed in the literature, for example [3, 12, 33]. Most of these are training based algorithms, though a few blind estimation techniques exist. Although it is not necessary to assume that $h(\delta) = 1$*, we do so for notational simplicity. Further, when the effects of error propagation are ignored during the FBF tap optimization, the FBF taps exactly cancel the post-cursor component of ISI.*

In an uncoded system the soft inputs $y(n)$ are fed into a hard-limiter and the output hard decisions $\hat{s}(n)$ form the estimates of the transmitted symbols, whereas in a coded system the soft inputs $y(n)$ are fed into the error correcting decoder for decoding and $d(n)$ are the tentative soft decisions used as inputs to the feedback filter to cancel ISI. Later, in Chapter 6 we will consider using the outputs of the error correcting decoder as the input to the feedback filter.

2.3 MSE Optimization of the Decision Box

The hard limiter of a conventional DFE is replaced by a memoryless soft decision device. Ideally, we would like to optimize the soft decision device to minimize the effects of error propagation. Though, typically BER is the metric of interest, the packet error rate (PER) would be a suitable performance metric in a packetized transmission environment. Hence, the choice of metric depends on the particular application.

In a coded scenario, it is almost impossible to analytically characterize the effect of error propagation on the error rate at the decoder output. However, if the MSE at the input of the decoder is minimized one can expect an improvement in the error rate at the decoder output. With this in mind an MSE optimized decision box is considered in this section.

Let $\Psi(\cdot)$ define the input/output relation of this soft decision device. This input/output relation is optimized by minimizing the mean squared error (MSE) cost function defined at the input of the decision device. The optimization problem can be formulated as,

$$\Psi_{opt} = \arg\min_{\Psi \in \mathcal{F}} \mathbf{E}[|y(n) - s(n)|^2] \tag{2.3}$$

where $\mathcal{F}$ is the space of one-to-one functions defined over $\Re$ and $E[(|y(n) - s(n)|)^2]$ is the MSE at the input of the decision device. For a given soft input $y(n)$, the decision $d(n)$ is

$$d(n) = \Psi\{y(n)\} \tag{2.4}$$

Let us now assume that the input to the decision device at the n^{th} time instant is $y(n)$ and the corresponding soft decision is $d(n)$. This soft decision is used to cancel the post-cursor ISI components of the received signal. Hence, this decision would contribute to the residual post-cursor ISI until the $n + L_b^{th}$ time instant. Here, the optimization problem is formulated as finding the optimal soft decision $d(n)$ that would minimize the expected value of squared error at the input of the decision device at the $n + 1^{th}$ time instant. Later in this section, we will see that this choice of $n + 1$ does not have any effect on the optimal decision device. The MSE at the input of the decision box for the $n + 1^{th}$ time instant is ,

$$\mathbf{J}(d_n, y_n) = \sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n)\mathbf{E}[(|y_{n+1} - s_{n+1}|)^2|s_n = S, y_n] \tag{2.5}$$

where $\Pr(s_n = S|y_n)$[1] is the conditional probability that the symbol S was transmitted. $\mathbf{E}[(|y_{n+1} - s_{n+1}|)^2 | s_n = S, y_n]$ is the expected value of the squared error for the $n+1$ th time instant, provided the soft decision at the n th time instant is $d(n)$ and $\mathbf{E}[\cdot]$ is the expectation operation. Now, $y(n+1)$ can be written as

$$y(n+1) = s(n+1) + b(1)[s(n) - d(n)] + z(n+1) \tag{2.6}$$

where $z(n)$ comprises the modeled post-cursor ISI, filtered noise, residual precursor ISI and residual post-cursor ISI. $z(n)$ can be written as,

$$\begin{aligned} z(n) = &\sum_{i=2}^{L_b} b(i)[s(n-i) - d(n-i)] + \sum_{i=0}^{L_f} f(i)w(n-i) \\ &+ \sum_{i=0}^{i=\delta-1} h(i)s(n+\delta-i) + \sum_{i=\delta+L_b+1}^{L_h} h(i)s(n+\delta-i) \end{aligned} \tag{2.7}$$

- **Claim 1:** $z(\cdot)$ is a zero mean random process.

  $w(n)$ is a zero-mean additive white Gaussian process. Hence the filtered noise is zero mean. If the constellation set $\mathcal{A}$ is symmetric about the origin, the residual ISI will be zero mean. Further, if we assume that the decision device is symmetrical about the origin, the modeled post-cursor ISI term is also zero mean. As we shall see later, the decision device is symmetrical for a symmetric constellation

- **Assumption 3:** $z(n+1)$ is conditionally uncorrelated to $[s(n) - d(n)|y(n)]$

  Only $w(n+1)$ component of the filtered noise is uncorrelated with the decision error $s(n) - d(n)$. The residual ISI components of $z(n+1)$ has some correlation with the decision error $[s(n) - d(n)]$, but we assume that the correlation is small enough to be ignored. This assumption simplifies the derivation of the soft decision device. Figure 2.2 compares the relative magnitudes of the

[1] $s(n)$ and s_n are used interchangeably.

auto-correlation of $z(n)$ and the weighted decision error $b(1)\{s(n) - d(n)\}$ with its cross-correlation, for a sample channel-feed forward combination of $h = [1\ 0.6\ \ -0.55\ \ -0.1]$. Note that, since the cross-correlation conditioned on $y(n)$ is very difficult to evaluate the unconditioned cross-correlation plot is shown. This gives some insight into the validity of the assumption.

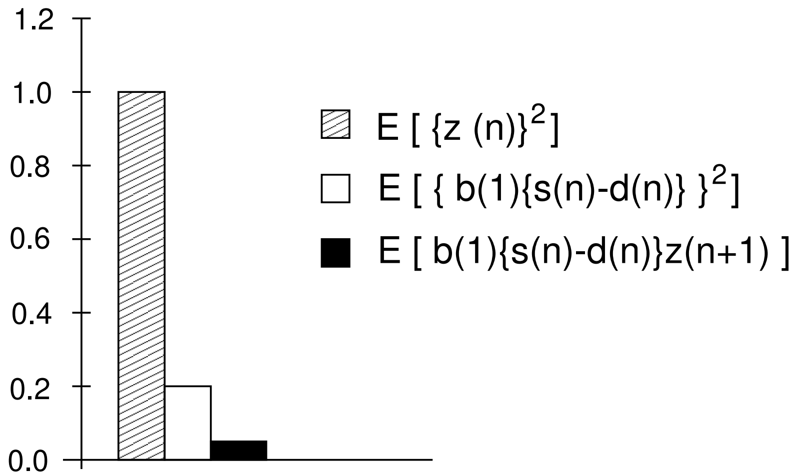

Figure 2.2: Noise and error correlations for h=[1 0.6 -0.55 -0.1]

Since, $d(n)$ has a one-to-one relationship with $y(n)$, conditioning on $d(n)$ is equivalent to conditioning on $y(n)$. Hence,

$$\begin{aligned}
\mathbf{E}[(|y_{n+1} - s_{n+1}|)^2|s_n = S, d_n] &= \mathbf{E}[\{|b(1)\{s(n) - d(n)\} + z(n+1)|\}^2|s(n) = S, d_n] \\
&= |b(1)\{S - d(n)\}|^2 + \mathbf{E}[\{|z(n+1)|\}^2|s_n = S, y_n] \\
&\quad + \mathbf{E}[b(1)\{S - d(n)\}z^*(n+1)|y_n] \\
&\quad + \mathbf{E}[b^*(1)\{S^* - d^*(n)\}z(n+1)(y_n] \qquad (2.8)
\end{aligned}$$

and since the last two terms go to zero,

$$\mathbf{E}[(|y_{n+1} - s_{n+1}|)^2|s_n = S] = \{|b(1)\{S - d(n)\}|\}^2 + \mathbf{E}[\{|z(n+1)|\}^2|s_n = S, y_n]$$

Note that the last term is independent of d_n. Hence equation (2.5) reduces to

$$\mathbf{J}(d_n, y_n) = \sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n)[|b(1)\{S - d(n)\}|\}^2 + \mathbf{E}[|z(n+1)|^2|s_n = S, y_n]]$$

Minimizing the MSE cost function,

$$\begin{aligned} & \tfrac{\partial \mathbf{J}(d_n)}{\partial d_n} = 0 \\ \Rightarrow \quad & \textstyle\sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n)[S - d_n] = 0 \\ \Rightarrow \quad & d_n = \textstyle\sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n) S \end{aligned} \tag{2.9}$$

Hence, from equation (2.9) the optimal decision device is given by,

$$\Psi_{opt}(y_n) = \mathbf{E}[s_n|y_n] \tag{2.10}$$

which is also the MMSE estimator of the transmitted symbol.

The input to the decision box can be written as

$$y(n) = s(n) + noise$$

where noise includes both residual ISI and AWGN. From equation (2.6),

$$noise = z(n) + b(1)\{s(n-1) - d(n-1)\}$$

- **Assumption 4:** *noise* is zero mean random process and has a complex Gaussian distribution

  From Claim 1 we can conclude that noise is zero mean. If the residual ISI has a reasonably large number of small terms then central limit theorem can be applied and a Gaussian approximation is valid. Figure 2.3 compares the distribution of noise at the input of the soft DFE decision box with a Gaussian of the same variance for a sample channel-feed forward combination of $h =$

$[1\ 0.6\ -0.55\ -0.1]$ *and an SNR of 8 dB. The value of the covariance matrix for this complex noise process was,*

$$\mathbf{K} = \begin{bmatrix} 0.2019 & 0.0020 \\ 0.0020 & 0.2024 \end{bmatrix}$$

These results seem to validate the above assumption.

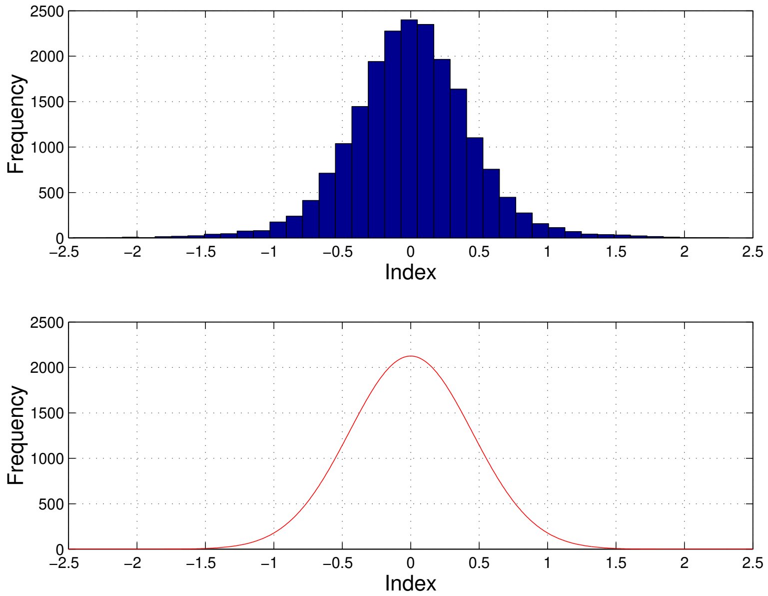

Figure 2.3: Distribution of *noise* for a soft DFE

Let *noise* have a total variance of σ^2. Under the above mentioned assumption, the decision device relationship for any linear two-dimensional constellation with equi-probable symbols is,

$$\Psi_{opt}(y_n) = \frac{\sum_{S\in\mathcal{A}} S \exp\left[\frac{-S^*S+y_n^*S+S^*y_n}{\sigma^2}\right]}{\sum_{S\in\mathcal{A}} \exp\left[\frac{-S^*S+y_n^*S+S^*y_n}{\sigma^2}\right]} \tag{2.11}$$

QPSK

For a QPSK modulation scheme the signal set contains the elements

$$\mathcal{A} = \left[\left(\frac{1}{\sqrt{2}} + \frac{j}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}} + \frac{j}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}} - \frac{j}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}} - \frac{j}{\sqrt{2}}\right)\right]$$

Since,

$$\Pr\left\{s_n = \left(\frac{1}{\sqrt{2}} + \frac{j}{\sqrt{2}}\right)\right\} + \Pr\left\{s_n = \left(\frac{1}{\sqrt{2}} - \frac{j}{\sqrt{2}}\right)\right\} = \Pr\left\{\mathrm{Re}\{s_n\} = \frac{1}{\sqrt{2}}\right\}$$

the real and imaginary parts of $\Psi_{opt}(y_n)$ can be decoupled as,

$$\mathrm{Re}\{\Psi_{opt}(y_n)\} = \frac{1}{\sqrt{2}}\left[\Pr\left\{\mathrm{Re}(s_n) = \frac{1}{\sqrt{2}} \,\Big|\, y_n\right\} - \Pr\left\{\mathrm{Re}(s_n) = \frac{-1}{\sqrt{2}}\right\} \,\Big|\, y_n\right]$$

$$\mathrm{Im}\{\Psi_{opt}(y_n)\} = \frac{1}{\sqrt{2}}\left[\Pr\left\{\mathrm{Im}(s_n) = \frac{1}{\sqrt{2}} \,\Big|\, y_n\right\} - \Pr\left\{\mathrm{Im}(s_n) = \frac{-1}{\sqrt{2}}\right\} \,\Big|\, y_n\right]$$

Hence the optimal[2] soft decision device is,

$$\mathrm{Re}\{\Psi_{opt}(y_n)\} = \frac{1}{\sqrt{2}}\tanh\left(\frac{\sqrt{2}\mathrm{Re}\{y(n)\}}{\sigma^2}\right) \tag{2.12}$$

$$\mathrm{Im}\{\Psi_{opt}(y_n)\} = \frac{1}{\sqrt{2}}\tanh\left(\frac{\sqrt{2}\mathrm{Im}\{y(n)\}}{\sigma^2}\right) \tag{2.13}$$

The soft decision device non-linearity is shown in Figure 2.4 for SNRs of 3, 8 and 13 dB. As the noise variance decreases the decision device tends more and more towards the hard-limiter. This is to be expected, since at very high SNRs error propagation will be absent and hard-limiter is optimal in that scenario.

8 PSK

The decision device non-linearity can be similarly obtained for an 8-PSK modulation scheme. The 8-PSK constellation is shown in Figure 2.5 where $a = \sin(\frac{\pi}{8})$

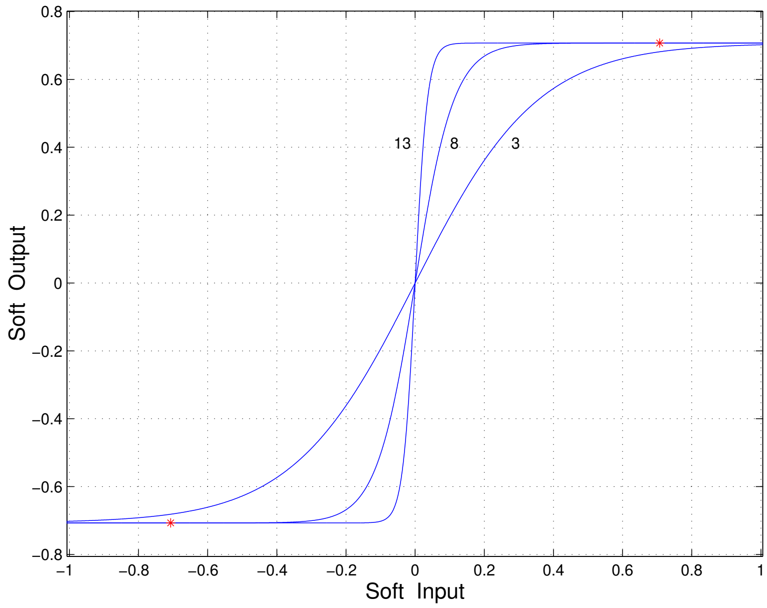

Figure 2.4: Soft DFE non-linearity for a QPSK modulation scheme

and $b = \cos(\frac{\pi}{8})$. Unlike the QPSK scheme, the real and imaginary parts of the decision $d(n)$ cannot be decoupled. The optimal decision device non-linearity is,

$$\mathrm{Re}\{d_n\} = \frac{a\sinh\left(\frac{2a\mathrm{Re}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2b\mathrm{Im}\{y_n\}}{\sigma^2}\right) + b\sinh\left(\frac{2b\mathrm{Re}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2a\mathrm{Im}\{y_n\}}{\sigma^2}\right)}{\cosh\left(\frac{2a\mathrm{Re}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2b\mathrm{Im}\{y_n\}}{\sigma^2}\right) + \cosh\left(\frac{2b\mathrm{Re}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2a\mathrm{Im}\{y_n\}}{\sigma^2}\right)}$$

and

$$\mathrm{Im}\{d_n\} = \frac{a\sinh\left(\frac{2a\mathrm{Im}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2b\mathrm{Re}\{y_n\}}{\sigma^2}\right) + b\sinh\left(\frac{2b\mathrm{Im}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2a\mathrm{Re}\{y_n\}}{\sigma^2}\right)}{\cosh\left(\frac{2a\mathrm{Im}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2b\mathrm{Re}\{y_n\}}{\sigma^2}\right) + \cosh\left(\frac{2b\mathrm{Im}\{y_n\}}{\sigma^2}\right)\cosh\left(\frac{2a\mathrm{Re}\{y_n\}}{\sigma^2}\right)}$$

Figure 2.6 shows the real part of the soft output of the decision device for an SNR value of 6 dB. The 3-D plot for the imaginary part of the soft output of the decision device is shown in Figure 2.7. This looks similar to the plot of the real part of soft output, but with the x and y axes interchanged.

[2]This has been independently proposed by Ariyavisitakul and Li in [36].

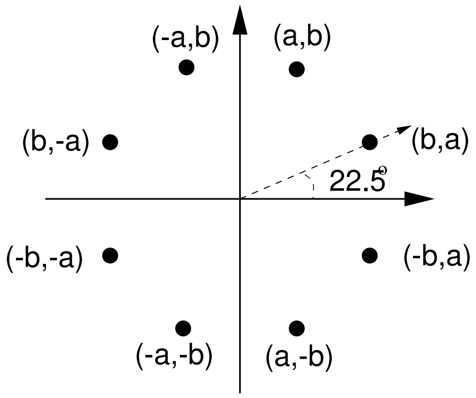

Figure 2.5: Constellation mapping for an 8PSK modulation scheme

16 QAM

The points in the 16 QAM constellation take on real and imaginary values from the set $\{-3,-1,1,3\}$. As in QPSK, the real and imaginary parts of the decision mapping $d(n)$ can be decoupled. The optimal decision box mapping is,

$$\mathrm{Re}\{\Psi_{opt}(y_n)\} = \frac{\sinh\left(\frac{2\mathrm{Re}\{y(n)\}}{\sigma^2}\right) + 3\exp(\frac{-8}{\sigma^2})\sinh\left(\frac{6\mathrm{Re}\{y(n)\}}{\sigma^2}\right)}{\cosh\left(\frac{2\mathrm{Re}\{y(n)\}}{\sigma^2}\right) + \exp(\frac{-8}{\sigma^2})\cosh\left(\frac{6\mathrm{Re}\{y(n)\}}{\sigma^2}\right)} \tag{2.14}$$

and

$$\mathrm{Im}\{\Psi_{opt}(y_n)\} = \frac{\sinh\left(\frac{2\mathrm{Im}\{y(n)\}}{\sigma^2}\right) + 3\exp(\frac{-8}{\sigma^2})\sinh\left(\frac{6\mathrm{Im}\{y(n)\}}{\sigma^2}\right)}{\cosh\left(\frac{2\mathrm{Im}\{y(n)\}}{\sigma^2}\right) + \exp(\frac{-8}{\sigma^2})\cosh\left(\frac{6\mathrm{Im}\{y(n)\}}{\sigma^2}\right)} \tag{2.15}$$

A plot of the decision device non-linearity is shown in Figure 4.2 of Chapter 4 for SNRs of 10, 15 and 20 dB. As expected, the decision device tends towards the hard-limiter for higher values of SNR.

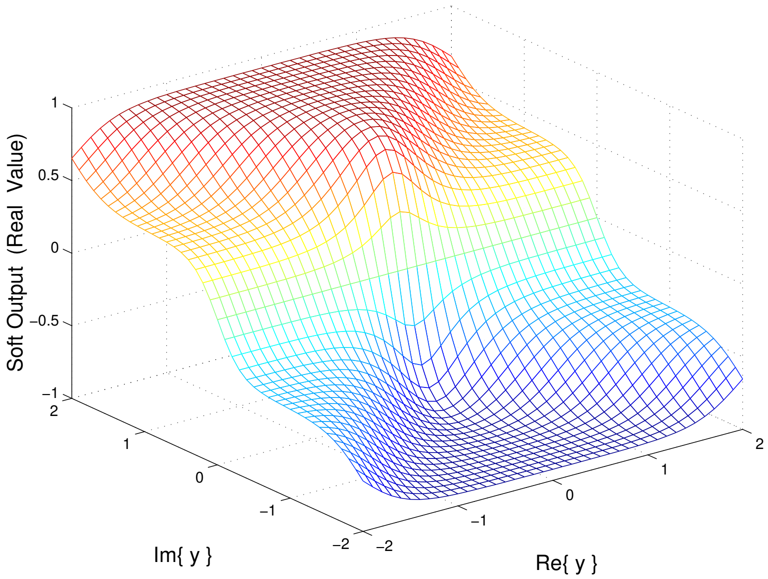

Figure 2.6: Soft decision device mapping for an 8-PSK modulation scheme : Real part

2.4 Simulation Results

Figure 2.8 compares the simulation based MSE performance curves of the soft DFE with a conventional DFE and an ideal DFE for a BPSK modulation scheme. A channel impulse response of C = [1 0.6 -0.55 -0.1] and a unit gain 1-tap feed-forward filter was used for this simulation. The channel is a mixed phase channel and its zeros are shown in Figure 2.9. Perfect knowledge of the optimal DFE coefficients was assumed at the receiver. The noise variance, σ^2, was estimated at the receiver using the update rule,

$$\sigma^2(n+1) = \lambda\sigma^2(n) + (1-\lambda)\{y(n) - \hat{s}(n)\}^2 \tag{2.16}$$

where λ is a decay rate slightly less than 1 for discounting past data. The value of λ is based on the degree of stationarity of the data. σ^2 is initialized with the noise

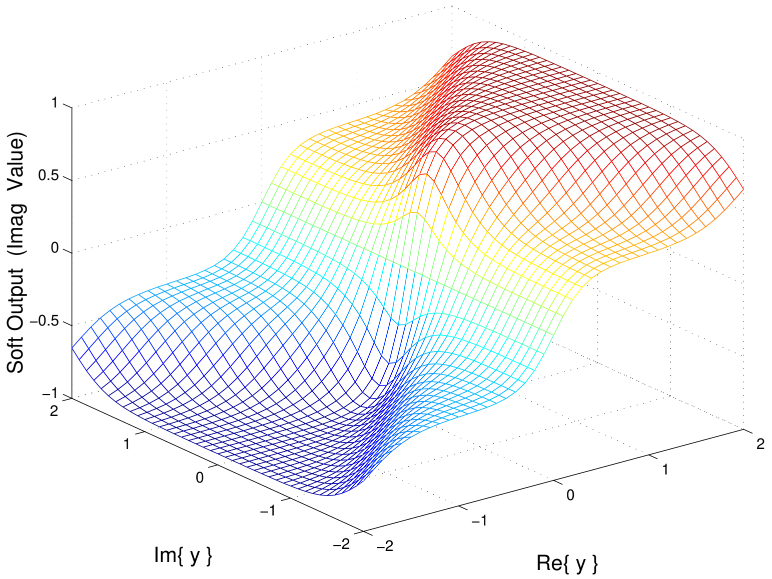

Figure 2.7: Soft decision device mapping for an 8-PSK modulation scheme : Imaginary part

variance of AWGN. In practical scenarios this initial value can be estimated during the training period. Performance results shows a reasonable error propagation mitigation for the soft DFE. At higher SNRs, primary errors (and hence the error bursts) occur less frequently and hence the MSE of the conventional DFE and the soft DFE are almost identical to that of the ideal DFE.

Figure 2.10 compares the simulation based MSE performance curves of the soft DFE with a conventional DFE and an ideal DFE for a 4-PAM modulation scheme. A channel impulse response of C = [1 0.05 0.6 -0.27 0.03 0.12 0.22] and a unit gain 1-tap feed-forward filter was used for this simulation. The channel is a minimum phase channel and its zeros are shown in Figure 2.11. Figure 2.12 compares the uncoded symbol error rate performance of the soft DFE over the conventional DFE for the above channel with a 4-PAM modulation scheme. We

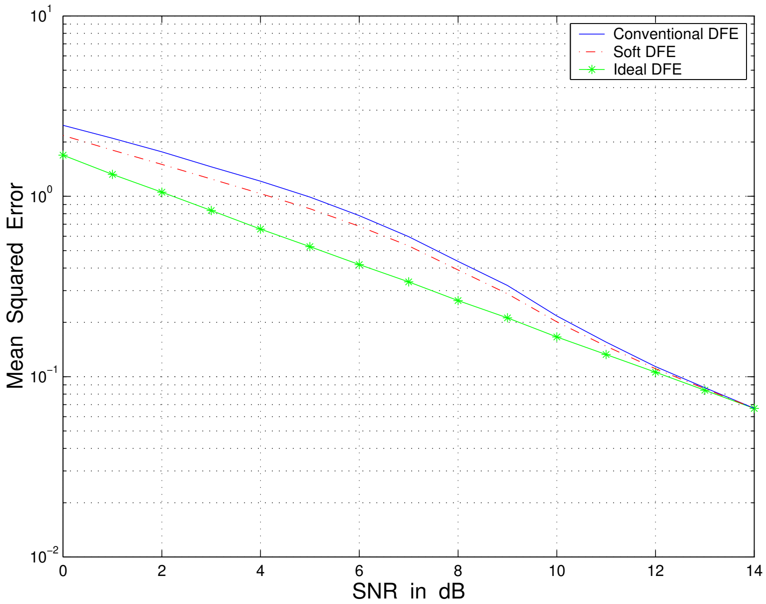

Figure 2.8: MSE performance comparison of soft DFE with conventional DFE

observe performance improvements for the soft DFE over the conventional DFE.

Figure 2.13 compares the performance of a soft DFE with a conventional DFE for a coded system for the channel C = [1 0.05 0.6 -0.27 0.03 0.12 0.22] and a 4-PAM modulation scheme. A rate (1/2) convolutional code with a constraint length of 3 was used as the error correcting code. The generator polynomials of the code were $g_1(D) = D^2 + 1$ and $g_2(D) = D^2 + D + 1$. A soft input Viterbi decoder with a decoding delay of atleast 12 symbols was used. At a BER of 10^{-2} the error propagation gap is about 5.4 dB and the soft DFE provides an improvement of about 1.6 dB over the conventional DFE. Note that the error propagation gap is larger for the coded system than that of the uncoded system. These simulation results demonstrate the advantage of using a soft DFE over the conventional DFE. The purpose of the above synthetic channel experiments is to demonstrate the effectiveness of the soft DFE. More elaborate simulations are presented in Chapter

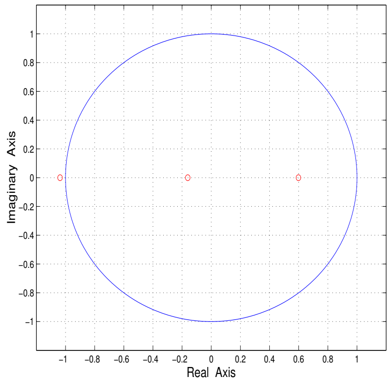

Figure 2.9: Channel zeros for C = [1   0.6   -0.55   -0.1]

5 for realistic channels.

2.5 BER Optimization of the Decision Box

Similarly, we can optimize the decision-box to minimize BER rather than MSE. Such a soft DFE would be of interest in an uncoded scenario or in a coded scenario when only a hard input channel decoder is used. In this section we derive the BER optimized decision box for the BPSK constellation. The BER optimization problem can be defined as,

$$\Psi_{opt} = \arg\min_{\Psi \in \mathcal{F}} \Pr[\text{sign}\{y(n)\} \neq s(n)] \tag{2.17}$$

where $\mathcal{F}$ is the space of one-to-one functions defined over $\Re$ and sign$\{y(n)\}$ is the bit decision[3]. The BER cost function is,

$$\mathbf{J}(d_n, y_n) = \sum_{S \in \{\pm 1\}} \Pr(s_n = S | y_n) \mathcal{P}(d_n, s_n = S) \tag{2.18}$$

[3] These are the symbol hard-decisions and are different from the tentative soft decisions used by the FBF for interference cancellation.

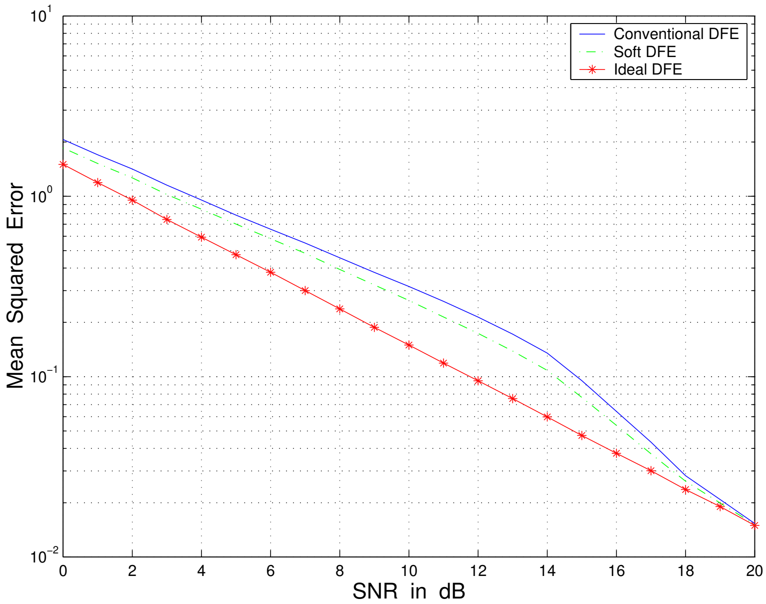

Figure 2.10: MSE performance comparison of soft DFE with conventional DFE

where $\mathcal{P}(d_n, s_n = S)$ gives the probability of error at the receiver for the $n+1^{\,th}$ time instant, given that the soft decision at the $n^{\,th}$ time instant is $d(n)$ and S is the transmitted signal at $n^{\,th}$ time instant.

$$\mathcal{P}(d_n, s_n = S) = \sum_{S' \in \{\pm 1\}} \Pr(s_{n+1} = S') \Pr[\text{sign}\{y(n+1)\} \neq S' | s_n = S, s_{n+1}S']$$

Recall equation (2.6)

$$y(n+1) = s(n+1) + b(1)[s(n) - d(n)] + z(n+1)$$

Assuming $z(n+1)$ to be a zero-mean complex Gaussian random process (similar to **Assumption 4**) with variance σ_z^2

$$\begin{aligned}\mathcal{P}(d_n, s_n = S) = \Pr(s_{n+1} = -1)\mathcal{Q}\left(\frac{1 - b(1)(s_n - d_n)}{\sigma_z}\right) \\ + \Pr(s_{n+1} = 1)\mathcal{Q}\left(\frac{1 + b(1)(s_n - d_n)}{\sigma_z}\right)\end{aligned} \tag{2.19}$$

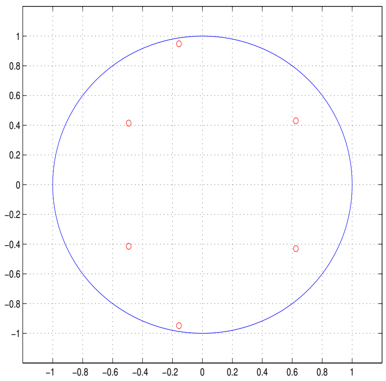

Figure 2.11: Channel zeros for C = [1 0.05 0.6 -0.27 0.03 0.12 0.22]

where $\mathcal{Q}(\cdot)$ is the Q-function defined as,

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(\frac{-Z^2}{2})dz \tag{2.20}$$

where $\mathcal{Q}(\cdot)$ is the Q-function defined as, If the transmitted symbols are equi-probable

$$\mathcal{P}(d_n, s_n = S) = \frac{1}{2}\left\{\mathcal{Q}\left(\frac{1 - b(1)(S - d_n)}{\sigma_z}\right) + \mathcal{Q}\left(\frac{1 + b(1)(S - d_n)}{\sigma_z}\right)\right\} \tag{2.21}$$

Now,

$$\begin{aligned}\mathbf{J}(d_n, y_n) = \frac{1}{2}\Bigg[&\Pr(s_n = 1|y_n)\left\{\mathcal{Q}\left(\frac{1 - b(1)(1 - d_n)}{\sigma_z}\right) + \mathcal{Q}\left(\frac{1 + b(1)(1 - d_n)}{\sigma_z}\right)\right\}\\ &+ \Pr(s_n = -1|y_n)\left\{\mathcal{Q}\left(\frac{1 + b(1)(1 + d_n)}{\sigma_z}\right) + \mathcal{Q}\left(\frac{1 - b(1)(1 + d_n)}{\sigma_z}\right)\right\}\Bigg]\end{aligned}$$

Minimizing the BER cost function[4],

$$\frac{\partial \mathbf{J}(d_n)}{\partial d_n} = 0 \tag{2.22}$$

The input-output relation of the decision device is given in implicit form as,

$$\exp\left(\frac{2y_n}{\sigma^2}\right) = \frac{\exp\left(\frac{-2b^2(1)d_n}{\sigma_z^2}\right)\sinh\left\{\frac{b(1)(1+d_n)}{\sigma_z^2}\right\}}{\sinh\left\{\frac{b(1)(1-d_n)}{\sigma_z^2}\right\}} \tag{2.23}$$

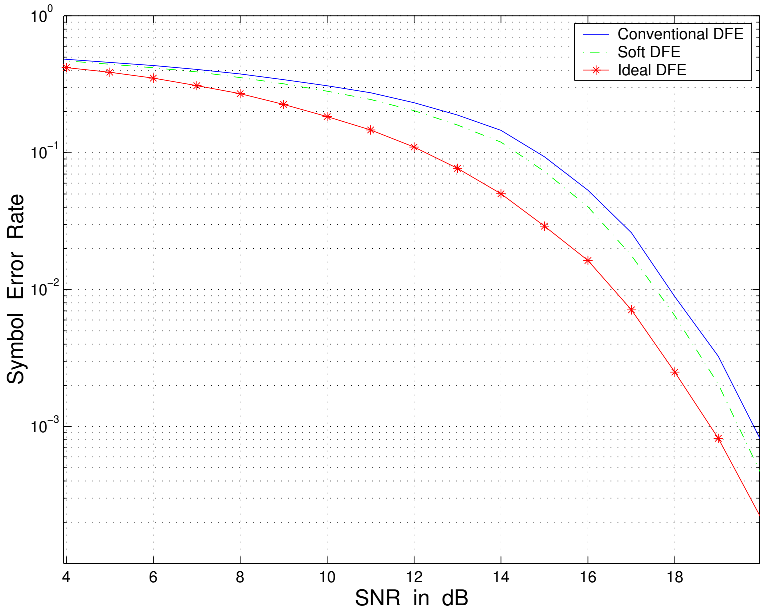

Figure 2.12: BER performance comparison of soft DFE for an uncoded system

In reality, it would not be possible to separately compute σ_z^2. Hence we might have to assume that σ_z^2 t σ^2. The non-linearity of the BER optimized decision box depends on the value of the feedback filter's coefficient $b(1)$. This is because we took into account only the effect of the first residual post-cursor ISI term and clumped all other terms with the noise component. Hence, for best results, we should formulate the problem by considering the residual post-cursor ISI term due to the most dominant feedback filter coefficient instead of $b(1)$. Figure 2.14 compares the BER optimized decision non-linearity with the MSE optimized decision non-linearity for $b(1) = 0.8$ and SNR values of 3, 8 and 13 dB. This dependence on the feedback-filter coefficient is a major disadvantage when we have adaptive scenarios, since the decision device has to be periodically updated. It is further complicated by

[4]See Appendix A for the details.

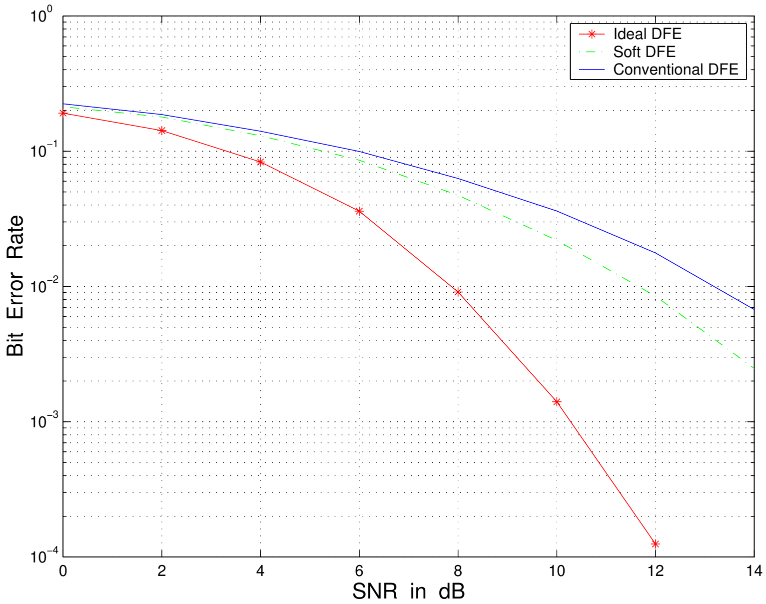

Figure 2.13: BER performance comparison of soft DFE for a coded system

the fact that the decision device non-linearity is available in implicit form. In such a case MSE optimization is a more viable option. One way to implement the BER optimized soft decision device would be to use a piece-wise linear approximation[5] of the decision device non-linearity.

Figure 2.15 compares the BER performance of a BER optimized soft DFE with a conventional DFE, ideal DFE and an MSE optimized soft DFE. The channel with impulse response C = [1   0.6   -0.55   -0.1] was chosen and a unit gain 1-tap feed-forward filter was used. Although, the BER optimized soft DFE performs better than the conventional DFE, its performance is almost the same as that of the MSE optimized soft DFE.

[5]See Chapter 4 for details

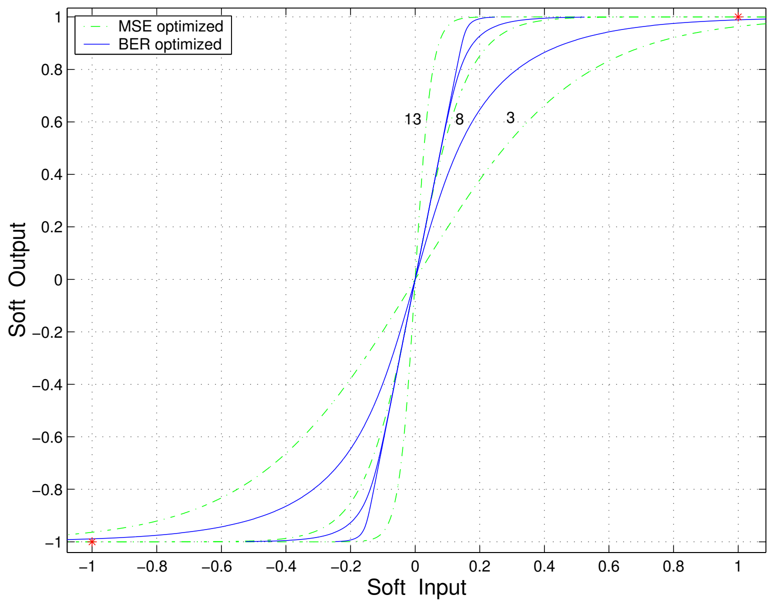

Figure 2.14: Comparison of BER optimized non-linearity with MSE optimized non-linearity

2.6 Gaussian Source

In this section we consider the joint MMSE optimization of the decision device non-linearity and DFE tap weights for a source that has a Gaussian distribution. The problem is formulated as,

$$[\mathbf{f}, \mathbf{b}, \Psi] = \arg\min_{\mathbf{f},\mathbf{b},\Psi} \mathbf{E}[|s(n) - y(n)|^2] \tag{2.24}$$

The source symbols $s(n)$ are assumed to be zero-mean, independent and identically distributed. Let $\mathbf{c}$ be the channel impulse response vector and $w(n)$ be the zero-mean additive white Gaussian noise. The received signal $r(n)$ can be written as,

$$r(n) = \mathbf{s}^T\mathbf{c} + w(n)$$

where $\mathbf{s}$ is the transmitted symbol vector.

The MMSE estimator that minimizes $\mathbf{E}[|s(n) - d(n)|^2]$ is $\mathbf{E}[\mathbf{s}|\mathbf{r}]$ and is given

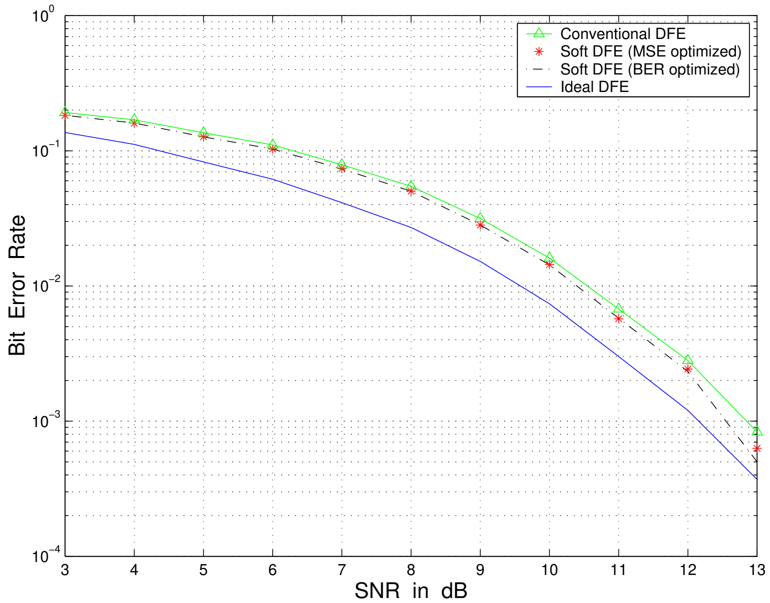

Figure 2.15: Performance comparison of BER optimized soft DFE

by,

$$\mathbf{d} = \mathbf{E}[\mathbf{s}\mathbf{r}^H](\mathbf{E}[\mathbf{r}\mathbf{r}^H])^{-1}\mathbf{r} \tag{2.25}$$

where $\mathbf{d}$ is the symbol decision vector. Notice that the MMSE estimator can be

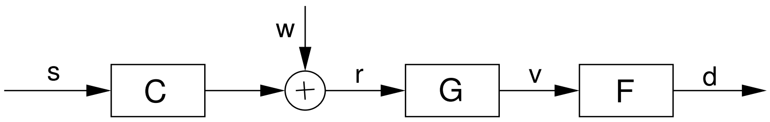

Figure 2.16: Block diagram of MMSE estimator for a Gaussian source

obtained by linear filtering the received vector $\mathbf{r}$. Let $C(Z)$ be the Z-transform of the channel impulse response, σ_s^2 be the source variance and σ_w^2 be the noise variance. Then, this linear filter can be causally implemented as shown in Figure 2.16, where $\mathbf{G}$ is the whitening filter. Now, from equation (2.25)

$$G(Z)G^*(\frac{1}{Z^*}) = \frac{1}{C(Z)C^*(\frac{1}{Z^*})\sigma_s^2 + \sigma_w^2} \tag{2.26}$$

We can solve for $G(Z)$ from the above equation and choose it to be minimum phase. The post-whitening filter, **F** has an impulse response,

$$f(k) = \mathbf{E}[s(n)v^*(n-k)] \qquad k = 0, 1, \cdots \tag{2.27}$$

When the channel is FIR, **G** can be implemented in auto-regressive form. Figure

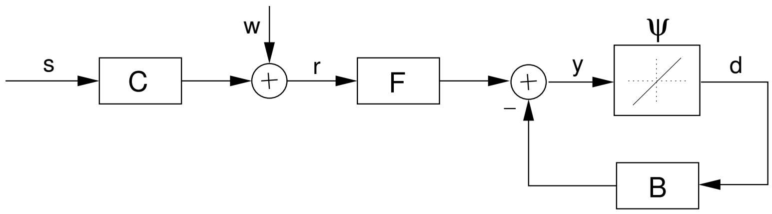

Figure 2.17: DFE structure for the MMSE estimator of a Gaussian source

2.17 shows the MMSE estimator implemented using a soft DFE structure. **F** is the feed-forward filter calculated using equation (2.27) and **B** is the feedback filter with,

$$B(Z) = 1 - G(Z) \tag{2.28}$$

The "non-linearity", Ψ, happens to be a straight line with a 45 degree slope. Note that due to the linearity of the decision device, $y(n) = \hat{s}(n)$ and hence,

$$\mathbf{E}[|s(n) - y(n)|^2] = \mathbf{E}[|s(n) - d(n)|^2] \tag{2.29}$$

Figure 2.18 compares the performance of the jointly optimized soft DFE with a conventional DFE for a 1024-PAM source with a Gaussian distribution. The ideal DFE bound is shown for the sake of comparison. The soft DFE outperforms the conventional DFE for a range of low SNR values. At very high SNR, the probability of decision errors are so low that the conventional DFE performs better. This discrepancy is due to the discrete nature of the transmitted source symbols.

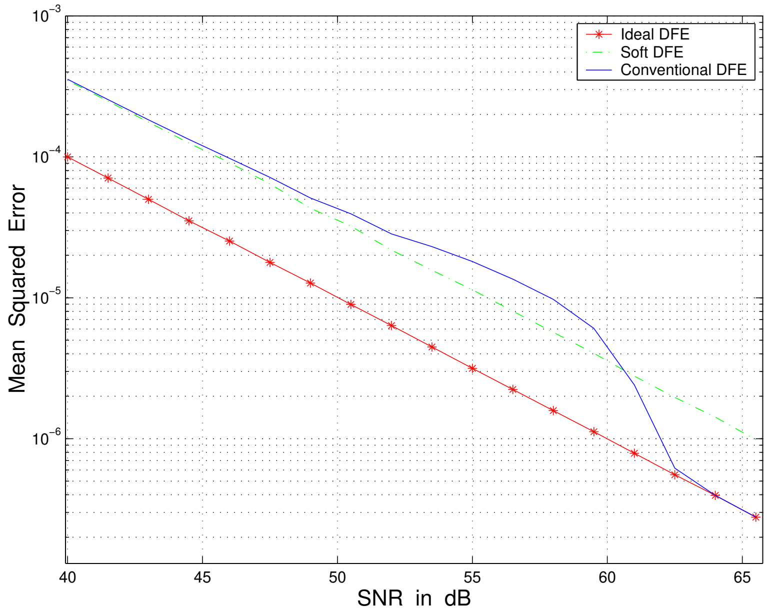

Figure 2.18: MSE performance for a 1024 PAM source with Gaussian distribution

2.7 Remarks

The error propagation gap between a conventional DFE and an ideal DFE depends on the modulation scheme used and the magnitude of the FBF taps. Typically, when the magnitude of the FBF taps increases the error propagation gap also increases. Figure 2.19 illustrates this increase in error propagation gap for the channel, C = [1 0.8 -0.6 0.3]. Compare this error propagation gap with Figure 2.8, which corresponds to the channel with smaller FBF taps, C = [1 0.6 -0.55 -0.1].

In this chapter we considered replacing the hard-limiter with a soft decision device with zero-order memory. We can alternately assume that the decision device has a first-order memory. The derivation of the MSE optimized decision device is almost identical to the zero-order memory case, but with an additional dependence

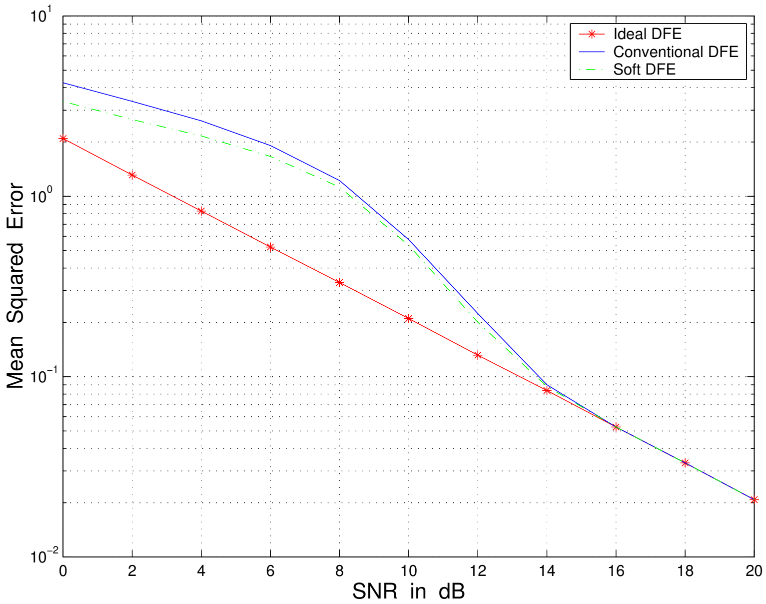

Figure 2.19: Error propagation gap for a channel with large FBF taps

of the decision d_n on y_{n-1}. The soft decision device relationship is given as

$$d_n = \mathbf{E}[s_n | y_n, y_{n-1}] \tag{2.30}$$

To evaluate d_n we need to evaluate $\Pr(s_n | y_n, y_{n-1})$. But,

$$\begin{aligned} \Pr(s_n | y_n, y_{n-1}) &= \frac{\Pr(s_n, y_n, y_{n-1})}{\Pr(y_n, y_{n-1})} \\ &= \frac{\Pr(y_n | y_{n-1}, s_n) \Pr(s_n)}{\Pr(y_n | y_{n-1})} \end{aligned} \tag{2.31}$$

Since the residual ISI component of s_{n-1} in y_n is unknown, it is very difficult to evaluate $\Pr(y_n | y_{n-1})$ and $\Pr(y_n | y_{n-1}, s_n)$.

Chapter 3

Performance Evaluation of Soft DFE

3.1 Modeling Error Propagation

In a conventional DFE, the quantizer output belongs to a finite alphabet set $\mathcal{A}$. The decision error can take only a finite number of values, M and for a feedback filter with L_b taps this results in M^{L_b} possible error patterns. The error dynamics of this system can be described by a state transition diagram. If the samples of noise at the input to the decision device are uncorrelated, then the probabilities of transition from one state to another depend only on those two states and the noise statistics. Error propagation, in a conventional DFE can be modeled as a Markov process, and the BER is then calculated from the state occupancy probabilities as determined by the state transition probability matrix [10].

Unfortunately, the outputs of the decision device in a soft DFE do not belong to a finite set and the decision error can take on real values in some interval. Hence, it is not possible to model the error dynamics of this system using a finite state

Markov process.

However, if the soft outputs of the decision device are quantized into K levels then the soft decision errors would be constrained to a finite number of values. As in the case of conventional DFE, this system, under the multi-level approximation, can be modeled using a finite state Markov process and the BER/MSE can be determined as before.

3.2 Constructing the Markov Chain

This section describes a method for constructing the Markov chain for the error dynamics of the multi-level quantization approximation of the soft DFE non-linearity. A BPSK constellation is assumed for the sake of brevity and clarity, although this method is applicable to any constellation. A $2K+1$ - level symmetric quantizer with mid-tread characteristic is used. The step size, Δ is defined as

$$\Delta = \frac{\sigma^2 \tanh^{-1}(0.99)}{K} \tag{3.1}$$

Figure 3.1 shows a 9 - level symmetric quantizer for the sigmoidal non-linearity corresponding to an SNR value of 3 dB for a BPSK modulation scheme. The quantization levels, α_i, for the decisions of the generic quantizer are given by,

$$\alpha_i = \begin{cases} \tanh\left(\frac{i\Delta}{\sigma^2}\right) & \text{if } i = 0, \pm 1, \ldots, \pm(K-1) \\ -1 & \text{if } i = -K \\ 1 & \text{if } i = K \end{cases} \tag{3.2}$$

The decision error, $e_k = s_k - \hat{s}_k$, can now take one of the following $4K+1$ possible values,

$$\beta_i = \begin{cases} 1 - \alpha_{K-i} & \text{if } i = 0, 1, \ldots, K \\ 1 + \alpha_{i-K} & \text{if } i = K+1, K+2, \ldots, 2K \end{cases} \tag{3.3}$$

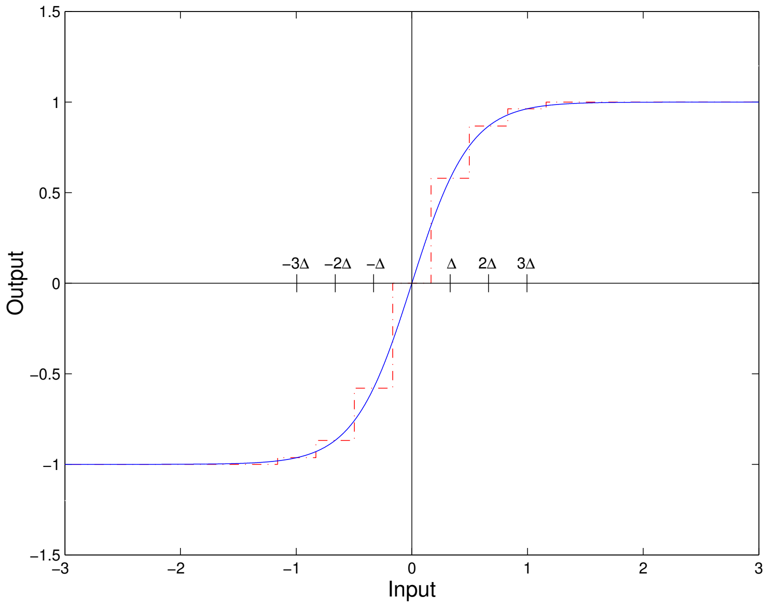

Figure 3.1: Multi-level quantization approximation of the soft DFE non-linearity

and

$$\beta_{-i} = -\beta_i$$

The feedback filter of the DFE has a L_b tap memory. Hence, the states of the state transition diagram can be uniquely specified by the error vector $(e_{k-1}, e_{k-2}, \ldots, e_{k-L_b})$. Since e_k can take one of $4K+1$ values, the corresponding Markov chain has $(4K+1)^{L_b}$ states. Let $(e_{k-1}, e_{k-2}, \ldots, e_{k-L_b})$ be the current state and e_k be the new error term. Then the error vector for the next state is $(e_k, e_{k-1}, \ldots, e_{k-L_b+1})$. When e_k is non-zero, this transition is possible only when the transmitted symbol "$s_k = \text{sign}(e_k)$". When e_k is zero, both "$s_k = 1$" and "$s_k = -1$" can give rise to this transition. The corresponding value of $\hat{s}_k$ is given by $\hat{s}_k = s_k - e_k$.

Figure 3.2 shows the state transition diagram for a $L_b = 1$ case with 3 - level quantization (i.e., $K = 1$). When $K = 1$, the possible values of $\hat{s}_k$ are $(-1, 0, 1)$ and hence e_k can take values from the finite set $(-2, -1, 0, 1, 2)$. $(s_k, \hat{s}_k)$ pairs that cause a particular transition are denoted on the edges of the state transition diagram.

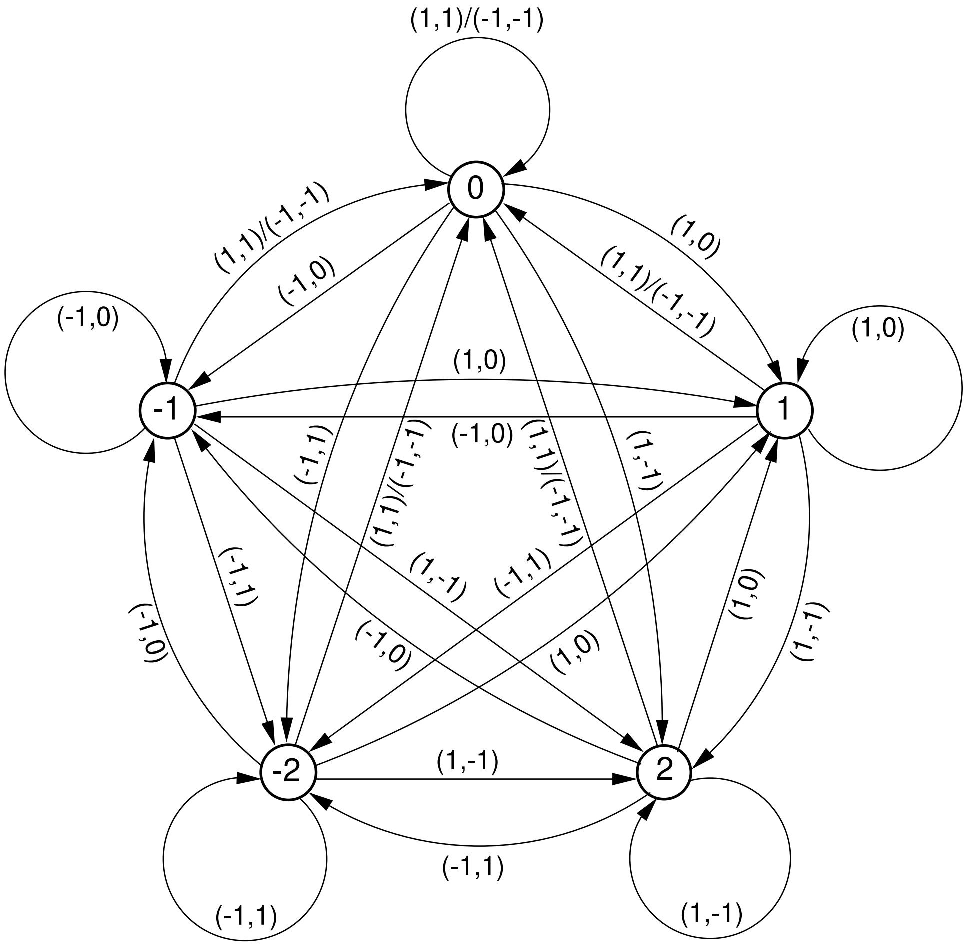

Figure 3.2: State transition diagram for soft DFE error propagation

Let y_k be the input to the decision box. Then

$$y_k = s_k + \mathbf{b}^T \mathbf{e} + z_k$$

where $\mathbf{b}$ is the feedback filter tap vector, $\mathbf{e}$ is the error vector corresponding to the current state and z_k is the noise term. Then,

$$\Pr\{\hat{s}_k = \alpha_i\} = \Pr\{a_i < y_k < b_i\} \tag{3.4}$$

where a_i and b_i are the left and right boundaries of the quantization interval that correspond to $\hat{s}_k = \alpha_i$.

$$a_i = \begin{cases} \frac{(2i-1)\Delta}{2} & \text{if } i \neq -K \\ -\infty & \text{if } i = -K \end{cases} \tag{3.5}$$

and

$$b_i = \begin{cases} \frac{(2i+1)\Delta}{2} & \text{if } i \neq K \\ \infty & \text{if } i = K \end{cases} \tag{3.6}$$

Hence,

$$\Pr\{\hat{s}_k = \alpha_i\} = \Pr\{a_i - s_k - \mathbf{b}^T\mathbf{e} < z_k < b_i - s_k - \mathbf{b}^T\mathbf{e}\} \tag{3.7}$$

The state transition probability is calculated using,

$$\begin{aligned} &\Pr\{e_k = \beta_i | e_{k-1}, \ldots, e_{k-L_b}\} \\ &\quad = \Pr\{a_i - \text{sign}(\beta_i) - \mathbf{b}^T\mathbf{e} < z_k < b_i - \text{sign}(\beta_i) - \mathbf{b}^T\mathbf{e}\} \Pr\{s_k = \text{sign}(\beta_i)\} \end{aligned} \tag{3.8}$$

when $i \neq 0$ and

$$\begin{aligned} \Pr\{e_k = \beta_i | e_{k-1}, \ldots, e_{k-L_b}\} = &\Pr\{a_i - 1 - \mathbf{b}^T\mathbf{e} < z_k < b_i - 1 - \mathbf{b}^T\mathbf{e}\} \Pr\{s_k = 1\} \\ &+ \Pr\{a_i + 1 - \mathbf{b}^T\mathbf{e} < z_k < b_i + 1 - \mathbf{b}^T\mathbf{e}\} \Pr\{s_k = -1\} \end{aligned}$$

when $i = 0$. From the state transition probabilities calculated using the above equations, the state transition matrix Γ can be constructed. The $(m,n)^{th}$ entry of the state transition matrix, Γ , is the transition probability from state $< m >$ to state $< n >$. The state occupancy probability vector μ satisfies,

$$\Gamma^T \mu = \mu \tag{3.9}$$

under the constraint

$$\sum_{i=1}^{(4K+1)^{L_b}} \mu_i = 1 \tag{3.10}$$

The solution to the above equation is obtained by finding the eigenvector of Γ^T with an eigenvalue of one and normalizing its sum to unity. The mean squared error for this error process is given by,

$$\begin{aligned} MSE &= E[(y_n - s_n)^2], \\ &= E[(z_n + \mathbf{b}^T\mathbf{e})^2] \end{aligned} \tag{3.11}$$

and since z_n is uncorrelated to the error vector $\mathbf{e}$

$$MSE = \sigma_z^2 + \sum_{i=1}^{(4K+1)^{L_b}} \mu_i \cdot (\mathbf{b}^T\mathbf{e}^{(i)})^2. \tag{3.12}$$

where $\mathbf{e}^{(i)}$ is the error vector corresponding to the $< i >^{th}$ state. We can alternately determine the bit error rate of the error process as,

$$BER = \sum_{i=1}^{(4K+1)^{L_b}} \frac{\mu_i \mathcal{H}\{\mathcal{I}(\mathbf{e}^{(i)})\}}{L_b} \tag{3.13}$$

where $\mathcal{I}(\mathbf{e}^{(i)})$ is the indicator function of the error vector $\mathbf{e}^{(i)}$ defined as,

$$\mathcal{I}(x) = \begin{cases} 0 & \text{if } -1 \leq x \leq 1 \\ 1 & \text{if } |x| > 1 \end{cases} \tag{3.14}$$

and $\mathcal{H}\{\mathcal{I}(\mathbf{e})\}$ gives the Hamming weight (i.e., the number of non-zero elements) of the indicator function of the error vector $\mathbf{e}$. The indicator function takes a non-zero value only when a particular value of soft error results in a hard-decision error.

3.3 Numerical Results

Figure 3.3 compares the theoretical performance results for a 3 - level and a 5 - level quantization approximation of the BPSK decision device non-linearity for the soft DFE with simulation results for the exact decision device non-linearity. The combined channel feed-forward impulse response was given by a 4 - tap transversal filter with coefficients [1   0.6   -0.55   -0.1]. We notice that the 5 - level approximation yields results that are as close as 0.1 dB. This justifies the use of the multi-level quantization approximation for performance evaluation of a soft DFE.

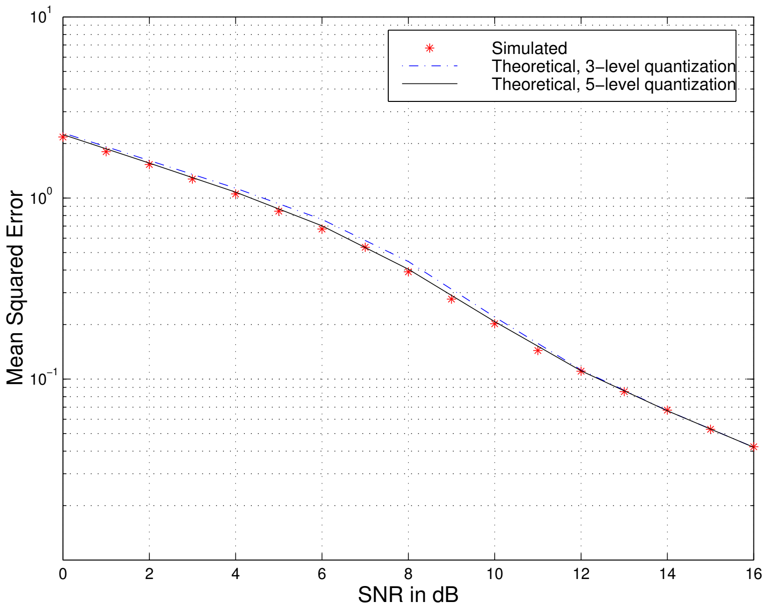

Figure 3.3: Performance results for h = [1   0.6   -0.55   -0.1]

Chapter 4

Implementation Issues

It is very difficult to implement the decision device nonlinearity of a soft DFE in a digital signal processor (DSP). This motivates the need for sub-optimal implementations of the decision device non-linearity. Two such sub-optimal techniques are discussed in this chapter.

4.1 Multi-Level Approximation

In the multi-level approximation, the soft decision nonlinearity is quantized into multiple levels. Hence, the output of the decision box can take only a finite number of values[1]. The domain of the input corresponding to each of the output values is an interval. This can be easily implemented by using a look-up-table to store the end-points[2] of each of the intervals and a comparator to compare the inputs with the end-points. Figure 3.1 of Chapter 3 depicts a 9-level symmetric quantizer for the BPSK case for an SNR of 3 dB. Exact theoretical results, based on the Markov model for error propagation, are shown in Figure 3.3 of Chapter 3. The

[1]Refer to equation (3.2).

[2]Refer to equations (3.5) and (3.6).

performance loss is very small (as low as 0.1 dB) for a reasonable choice of K.

4.2 Piece-wise Linear Approximation

Alternately, a simple piece-wise linear approximation for the soft decision non-linearity is possible. The implementation can be easily performed by scaling the input with a different factor for the various regions of the domain of the input that define the piece-wise linearity.

The points in the QPSK constellation assume real and imaginary values from the set $\left\{\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}$. Recall equation (2.12) and use the approximation,

$$\tanh(x)\ \mathrm{t} \begin{cases} x & \text{if } |x| < 1 \\ \operatorname{sign}(x) & \text{if } |x| \geq 1 \end{cases}$$

The piece-wise linear approximation for the QPSK case is,

$$\operatorname{Re}\{d(n)\} = \begin{cases} \frac{\operatorname{Re}\{y(n)\}}{\sigma^2} & \text{if } \frac{|\operatorname{Re}\{y(n)\}|}{\sigma^2} \leq \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\operatorname{sign}[\operatorname{Re}\{y(n)\}] & \text{if } \frac{|\operatorname{Re}\{y(n)\}|}{\sigma^2} > \frac{1}{\sqrt{2}} \end{cases} \tag{4.1}$$

$$\operatorname{Im}\{d(n)\} = \begin{cases} \frac{\operatorname{Im}\{y(n)\}}{\sigma^2} & \text{if } \frac{|\operatorname{Im}\{y(n)\}|}{\sigma^2} \leq \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\operatorname{sign}[\operatorname{Im}\{y(n)\}] & \text{if } \frac{|\operatorname{Im}\{y(n)\}|}{\sigma^2} > \frac{1}{\sqrt{2}} \end{cases} \tag{4.2}$$

Figure 4.1 shows the exact decision device mapping and the piece-wise linear approximations for SNRs of 3, 8 and 13 dB. For the 8-PSK case, it is very difficult to obtain a simple linear approximation due to the two-dimensional nature of the decision device input/output relation.

As in the QPSK case, the 16QAM soft decision mapping can be approximated piece-wise linearly. The piece-wise linear approximation for the real part is given

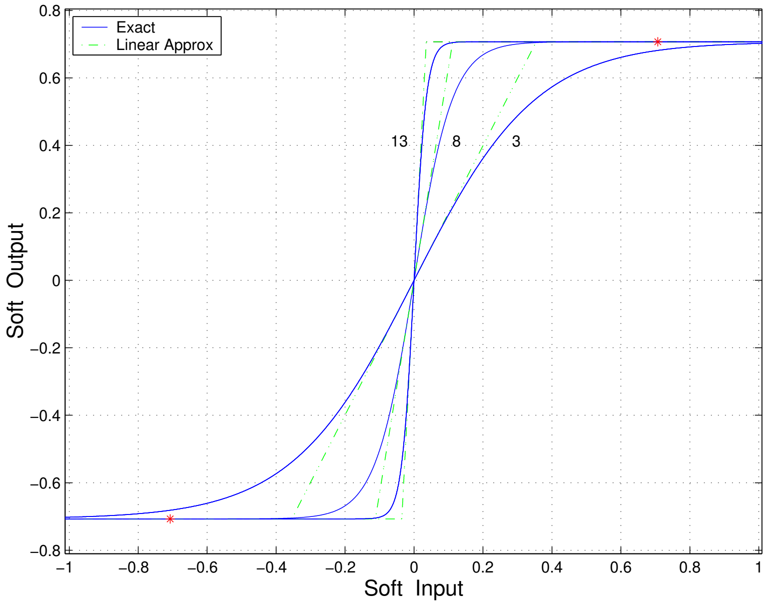

Figure 4.1: Piece-wise linear approximation of the soft DFE nonlinearity for a QPSK modulation

below. The imaginary part of the soft decision can be similarly determined.

$$\text{If } |\mathrm{Re}\{y_n\}| \le 1, \quad \mathrm{Re}\{d(n)\} = \begin{cases} \frac{2\mathrm{Re}\{y(n)\}}{\sigma^2} & \text{if } \frac{|2\mathrm{Re}\{y(n)\}|}{\sigma^2} \le 1 \\ \mathrm{sign}[\mathrm{Re}\{y(n)\}] & \text{if } \frac{|2\mathrm{Re}\{y(n)\}|}{\sigma^2} > 1 \end{cases}$$

$$\text{if } \mathrm{Re}\{y_n\} > 1, \quad \mathrm{Re}\{d(n)\} = \begin{cases} \frac{2(\mathrm{Re}\{y(n)\}-2)}{\sigma^2} + 2 & \text{if } \frac{|2(\mathrm{Re}\{y(n)\}-2)|}{\sigma^2} \le 1 \\ 1 & \text{if } \frac{2(\mathrm{Re}\{y(n)\}-2)}{\sigma^2} < -1 \\ 3 & \text{if } \frac{2(\mathrm{Re}\{y(n)\}-2)}{\sigma^2} > 1 \end{cases}$$

$$\text{if } \mathrm{Re}\{y_n\} < -1, \quad \mathrm{Re}\{d(n)\} = \begin{cases} \frac{2(\mathrm{Re}\{y(n)\}+2)}{\sigma^2} - 2 & \text{if } \frac{|2(\mathrm{Re}\{y(n)\}+2)|}{\sigma^2} \le 1 \\ -3 & \text{if } \frac{2(\mathrm{Re}\{y(n)\}+2)}{\sigma^2} < -1 \\ -1 & \text{if } \frac{2(\mathrm{Re}\{y(n)\}+2)}{\sigma^2} > 1 \end{cases}$$

The above approximation was obtained under the following assumption. The real part of the transmitted symbol takes one of four values $[\pm 1, \pm 3]$. For a given value of the decision device input $\mathrm{Re}(y_n)$, the two farthest symbols from $\mathrm{Re}(y_n)$

were assigned an aposteriori probability value of zero. Using this method, the decision device non-linearity was broken into three regions. For each of these regions, a piece-wise linear approximation was performed.

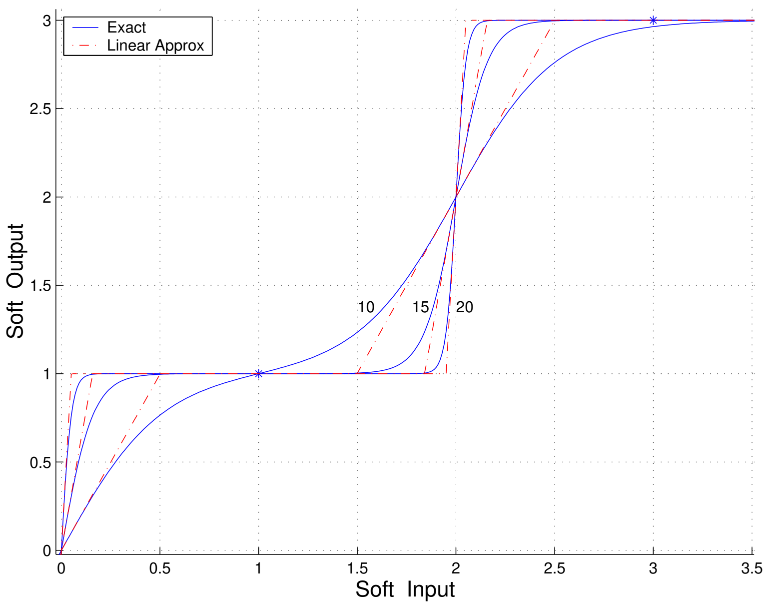

Figure 4.2: Piece-wise linear approximation of the soft DFE nonlinearity for a 16QAM modulation

Figure 4.2 shows the exact soft DFE mapping and the piece-wise linear approximation for SNRs of 10, 15 and 20 dB. Since the mapping is symmetric about the origin, only the positive half is shown.

BER Optimized Soft Decision Device

The piece-wise linear approximation can also be performed on the BER optimized soft decision device for BPSK as follows.

$$d(n) = \begin{cases} \frac{y(n)}{\gamma} & \text{if } \frac{|y(n)|}{\gamma} \leq 1 \\ \text{sign}[y(n)] & \text{if } \frac{|y(n)|}{\gamma} > 1 \end{cases} \tag{4.3}$$

where

$$\gamma = b(1)\coth(\frac{b(1)}{\sigma^2}) - b(1)^2 \tag{4.4}$$

is the gradient ($\frac{\partial d_n}{\partial y_n}$) of equation (2.23) evaluated at the origin.

4.3 Simulation Results

Figure 4.3 compares the MSE performance of the MSE optimized piece-wise linear approximation for the soft DFE decision device with the exact soft DFE. The simulations were performed for a 4-PAM modulation scheme and for a combined channel-feed forward impulse response of $h = [1\ 0.05\ 0.6\ -0.27\ 0.03\ 0.12\ 0.22]$. Exact knowledge of the feedback filter coefficients were assumed at the receiver. Noise variance was estimated as indicated in Section 2.4 of Chapter 2. Figure 4.4 compares the performance of a soft DFE with a MSE optimized piece-wise linear approximation with an exact soft DFE and conventional DFE for a coded system with a 4-PAM modulation scheme and the same sample channel. A rate (1/2) convolutional code with a constraint length of 3 was used as the error correcting code. The generator polynomials of the code were $g_1(D) = D^2 + 1$ and $g_2(D) = D^2 + D + 1$. A soft input Viterbi decoder with a decoding delay of atleast 12 symbols was used. At a BER of 10^{-2} the piece-wise linear approximation suffers only a loss of 0.4 dB over the exact soft DFE.

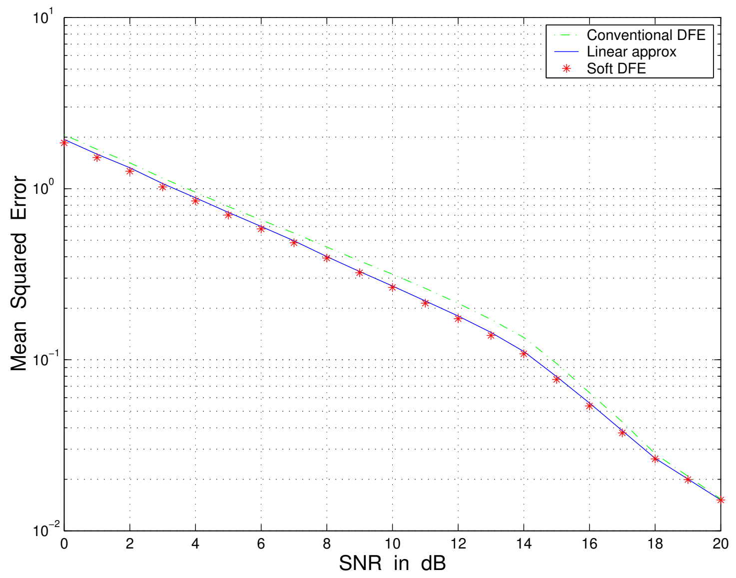

Figure 4.3: MSE performance comparison of soft DFE with a piece-wise linear approximation

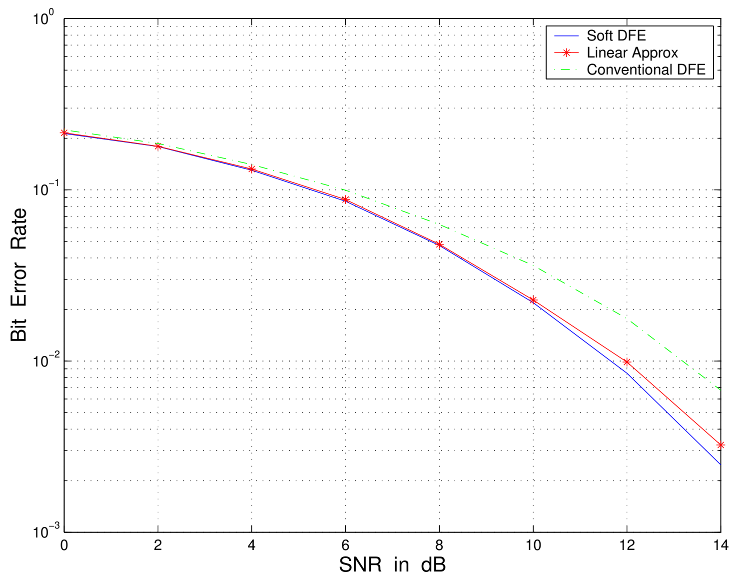

Figure 4.4: BER performance comparison of soft DFE with a piece-wise linear approximation for a coded system

Chapter 5

Simulation Examples

In this chapter simulation results on the performance of soft DFE for measured microwave and broadband wireless channels are presented.

5.1 Microwave Channels

Figure 5.1 shows the squared magnitude impulse response of an FIR model of a digital microwave radio channel. The complex-valued baseband channel response has been "fractionally sampled" at twice the baud interval (Baud Rate = 30 M-baud). This response was extracted from the processing of field measurements of received signals, courtesy of Applied Signal Technology, Sunnyvale, CA, available at the Signal Processing Information Base (SPIB). A DFE with 64 fractionally-spaced feed-forward taps and 150 feedback taps was used (for a delay spread of 182 symbols). Training was assumed to be perfect, i.e., optimum DFE tap gains (for FFF and FBF taps) were obtained numerically by solving the minimum mean-squared error (MMSE) equations. For the soft DFE approach we use the actual value of mmse (assuming no error propagation) after tap gain computations for

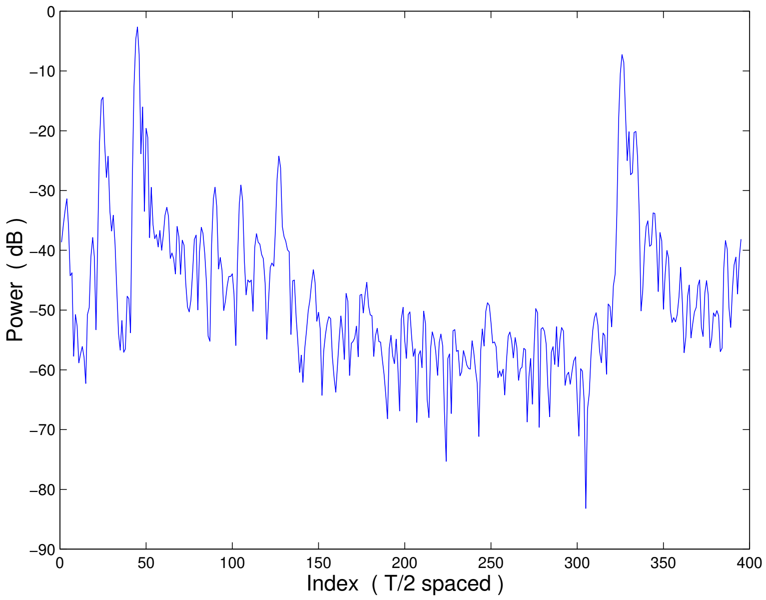

Figure 5.1: Magnitude response of a measured microwave channel with severe echoes

the noise variance. Figure 5.2 compares the MSE performance of the soft DFE with that of the conventional and ideal DFE for an uncoded QPSK modulation scheme. We observe improvements similar to that of the synthetic channel model simulation experiment (Figure 2.8) of Chapter 2.

Figure 5.3 compares the BER performance of the soft DFE with the conventional and ideal DFE for a convolutionally coded scheme with 4-PAM modulation. A (2,1,3) convolutional code with a generator matrix $[1 + D^2 \quad 1 + D + D^2]$ was used as the error correcting code.

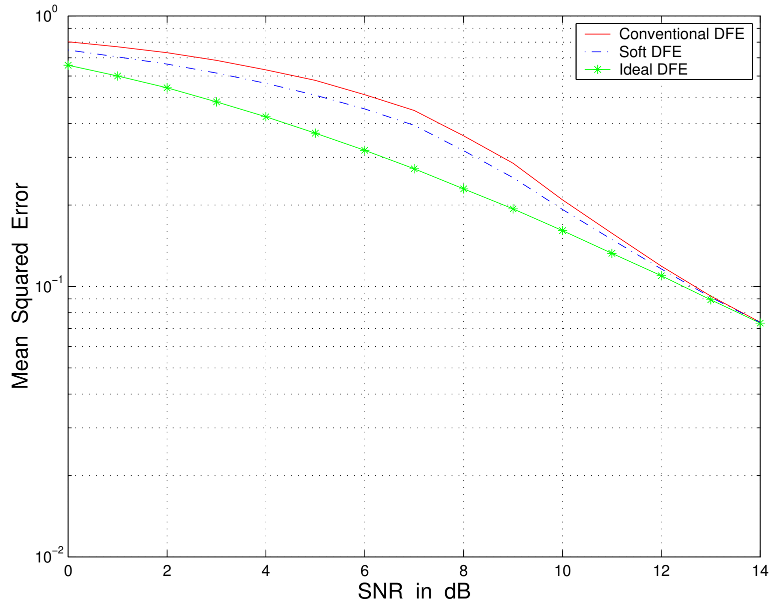

Figure 5.2: MSE performance comparison of soft DFE for a microwave channel model (uncoded scenario)

5.2 Broadband Wireless Channels

The multi-path delay profiles specified by GSM [41] (with slight modifications) for the urban environment and a hilly terrain environment were used. The power profiles for the environments are shown in Figures 5.4 and 5.5. All the paths in the above profiles were assumed to be independently Rayleigh fading. A channel bit rate of 2 Mbits/s (the symbol period is 1 μ second) was assumed. A square-root raised cosine pulse shaping filter with a 12.5% roll-off factor was used. A DFE with 8 fractionally-spaced feed-forward taps and 8 feedback taps was used for the urban environment, while a 32 tap fractionally-spaced FFF and 20 tap FBF was used for the hilly terrain environment. As before, training was assumed to be perfect, i.e., optimum DFE tap gains were obtained numerically by solving the minimum

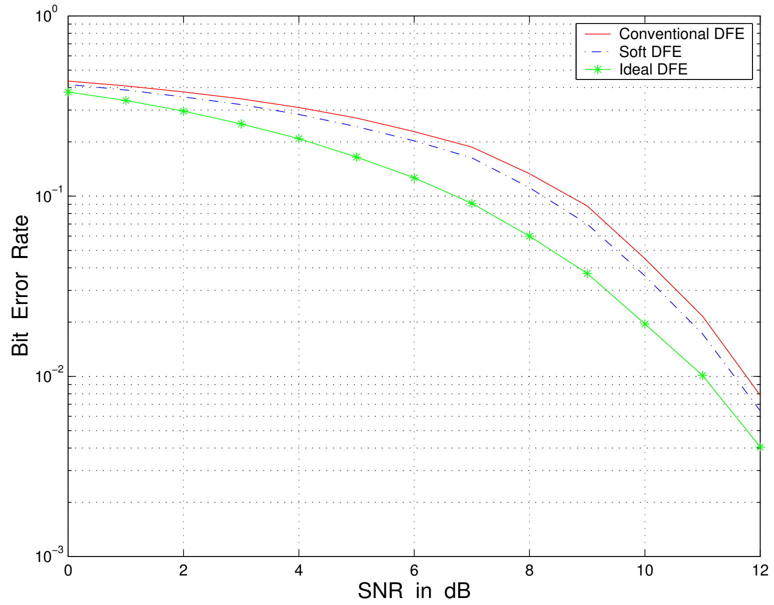

Figure 5.3: BER performance comparison of soft DFE for a microwave channel model (coded scenario)

mean-squared error (MMSE) equations. For the soft DFE approach we use the actual value of mmse (assuming no error propagation) after tap gain computations for the noise variance.

Simulations were performed for an ensemble of randomly generated channels with either the urban or the hilly terrain environment power profiles and averaging the results. Figure 5.6 shows the MSE comparison plot for an uncoded QPSK system for the urban environment. Figure 5.8 shows the BER performance curves for four randomly generated channels with the urban power profile. The BER plots were generated for a coded system with a 4-PAM modulation scheme. The convolutional code used was the same as that of the microwave channel experiment. Figure 5.7 compares the average BER performance averaged over an ensemble of urban channels.

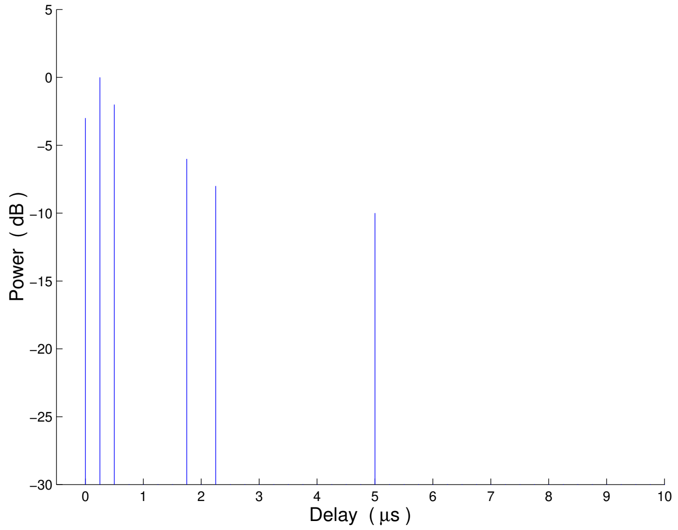

Figure 5.4: Power profile of channel for a typical urban environment

Figure 5.9 shows the MSE comparison plot for an uncoded QPSK system for the hilly terrain environment. Figure 5.11 shows the coded BER performance curves for four randomly generated channels with the hilly terrain power profile. Figure 5.7 compares the average BER performance averaged over an ensemble of hilly terrain channels. The performance of a particular channel depends on the noise gain of the FFF and the error propagation gap due to the FBF. Hence, at high SNRs the BER performance is limited by the BER of the bad channels. Hence we notice a performance floor for the averaged BER curve.

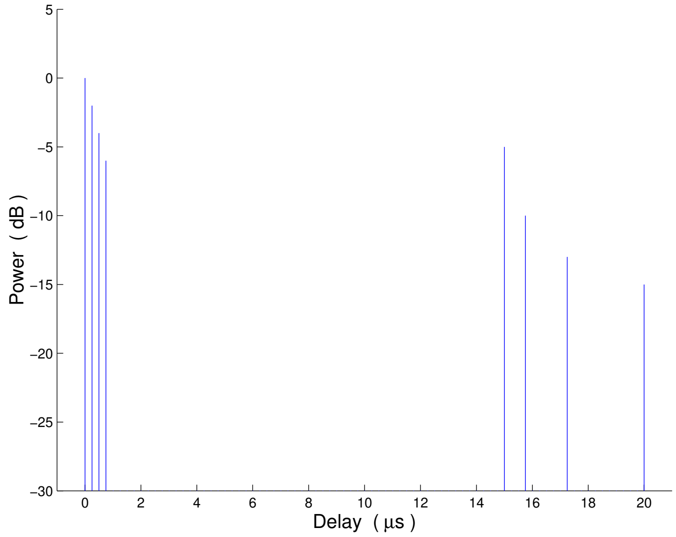

Figure 5.5: Power profile of channel for a hilly terrain environment

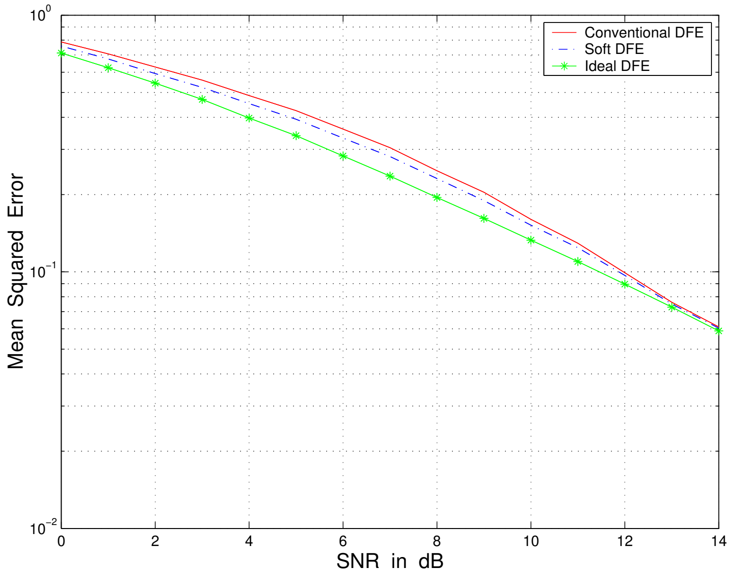

Figure 5.6: MSE performance comparison of soft DFE for an urban environment (uncoded scenario)

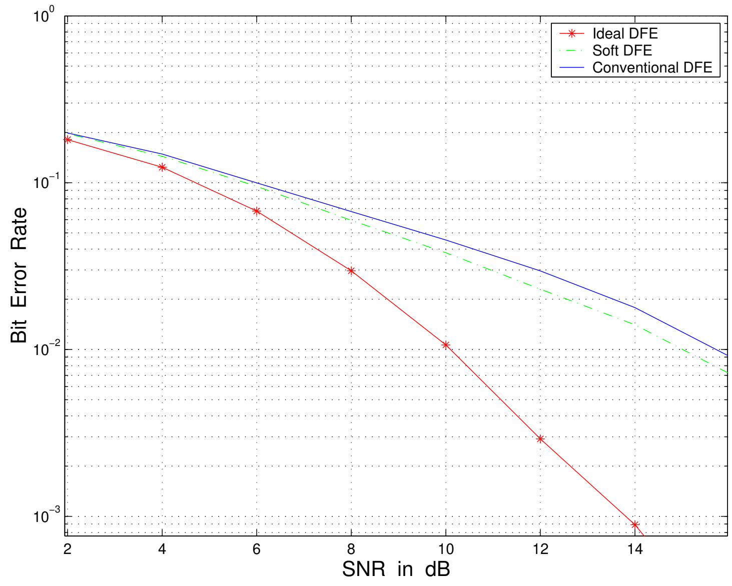

Figure 5.7: BER performance comparison of soft DFE for an urban environment (coded scenario)

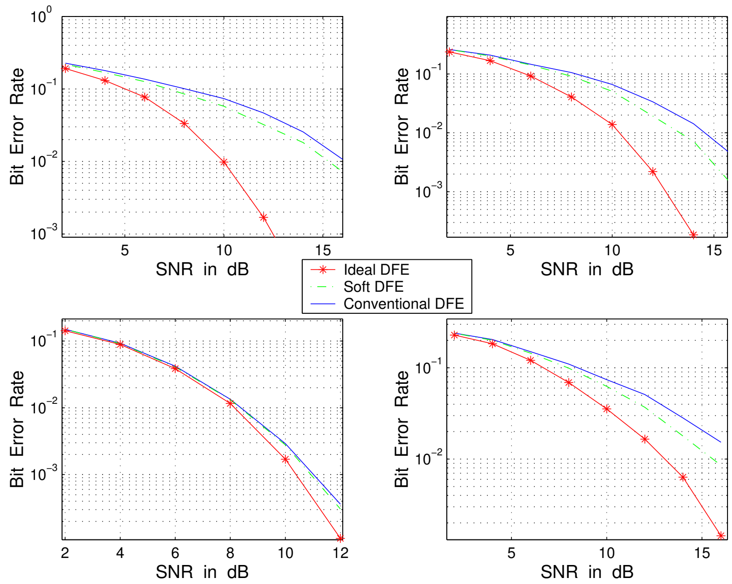

Figure 5.8: BER performance plots of soft DFE for a few urban channels (coded scenario)

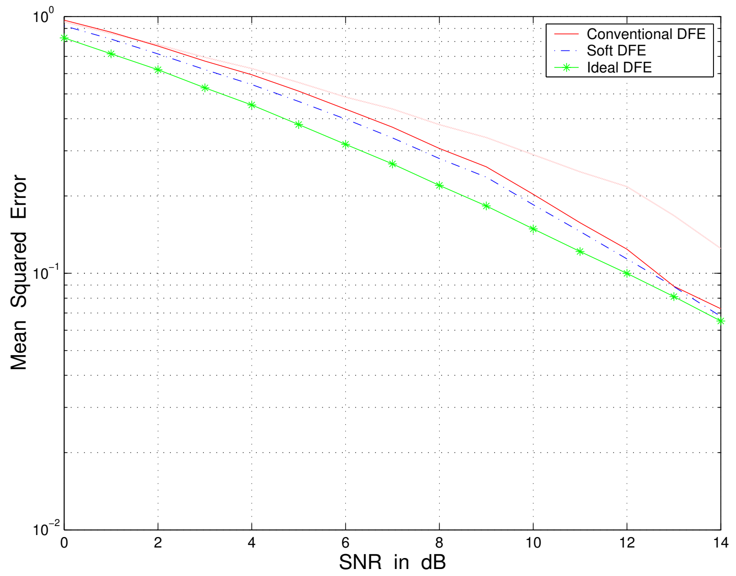

Figure 5.9: MSE performance comparison of soft DFE for a hilly terrain environment (uncoded scenario)

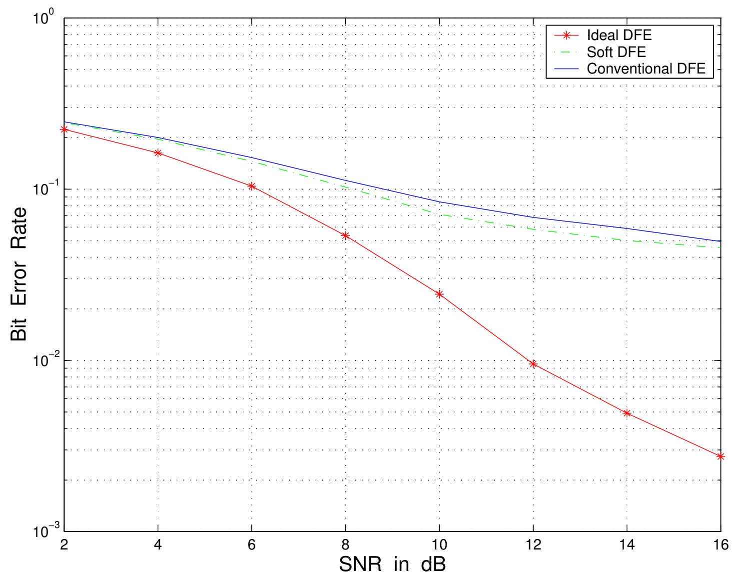

Figure 5.10: BER performance comparison of soft DFE for a hilly terrain environment (coded scenario)

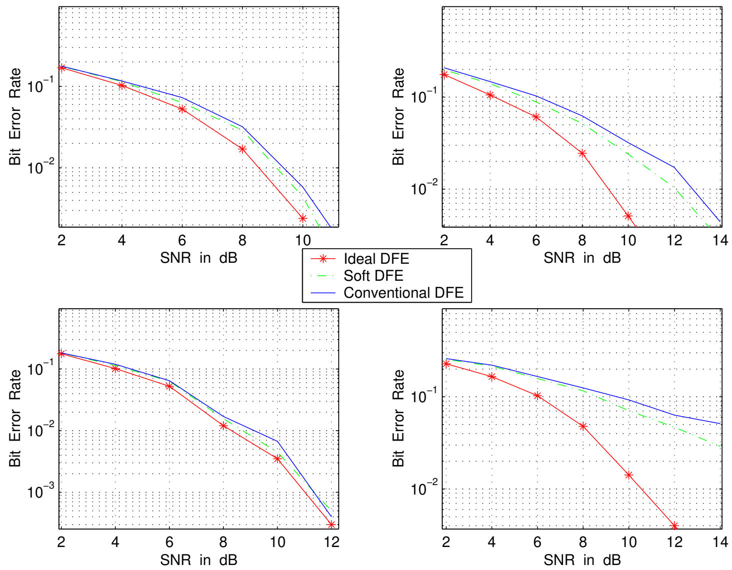

Figure 5.11: BER performance plots of soft DFE for a few hilly terrain channels (coded scenario)

Chapter 6

Turbo DFE

6.1 Joint Equalization and Coding

Error correction coding is used to reduce the error rate by introducing certain transmission redundancy. Most coding techniques are designed to correct random errors. However, in the presence of ISI the error-correction code suffers performance degradation[1]. In order to mitigate this undesirable interference, the channel needs to be equalized. The channel, due to its memory behaves like a convolutional encoder. Hence, improved performance can be achieved by combining the memory in the channel and the error correction code to construct a super trellis and use the maximum-likelihood technique implemented with the trellis-based Viterbi algorithm [7] for detection and decoding. However, the complexity of the Viterbi algorithm is proportional to the number of states in the trellis which grows exponentially with the combined channel/encoder memory length. When the channel has long memory, this approach becomes impractical.

[1]The probability of error expression for the Viterbi decoder in the presence of correlated noise is derived in [30].

Most approaches involve a sub-optimal equalization stage followed by decoding. A DFE would be a good candidate for the equalization stage. However, the error-propagation mechanism of the DFE makes it particularly difficult to successfully combine DFE with error-correction coding. In order to remedy this problem, several techniques have been proposed in the past, based on: (i) transmitter precoding [4, 6], (ii) use of interleaving to introduce delay necessary for reliable feedback [19, 22], (iii) use of tentative or soft decisions for feedback [29, 30, 36]. Transmitter precoding requires knowledge of the channel response by the transmitter, and thus might be inapplicable if the transmission is broadcast or if the channel is time-varying. The interleaving approach requires the DFE to be implemented in noise-predictive form, with independent adaptations of the feed forward and feedback filters. From a performance perspective, a noise predictive DFE is not recommended for a low-complexity DFE with a short feed-forward filter [19]. In addition to the above mentioned techniques, schemes that integrate the error detection code in to the DFE structure have also been proposed [31].

[28] proposes a turbo-equalization scheme, referred by analogy with turbo-codes, where the detection and decoding are done iteratively to improve performance. Simulation results indicate complete suppression of ISI for Gaussian channels. In this chapter, we propose a turbo-like scheme akin to that in [28] with a DFE as the symbol detector and a hard decision Viterbi algorithm for channel decoding.

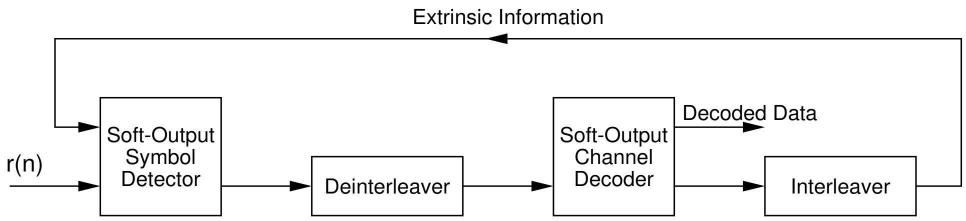

Figure 6.1: Principle of turbo-equalization (under zero internal delay assumption)

6.2 Turbo-Equalization

Figure 6.1 illustrates the general structure of a turbo-equalizer under the zero internal delay assumption. In order to use a soft-input channel decoder, the symbol detector has to provide information about the reliability of the symbols estimated. This information may be obtained by using the Soft-Output Viterbi Algorithm (SOVA) [21] that associates an estimation of the logarithm of its likelihood ratio (LLR) to each detected symbol. After deinterleaving, the SOVA decoder provides a new LLR value for the decoded symbol. This information may be used through a feedback loop by the symbol detector, after interleaving. This is the basis of the turbo-equalization principle. Unfortunately, the computational complexity of the SOVA detector increases exponentially with the length of the channel impulse response. In such a scenario a sub-optimal equalizer, like the DFE, can be used for the equalization block.

In the turbo-equalization principle, the information that is interchanged between the detector and decoder is the symbol reliability. This technique cannot be directly applied if the DFE is used as the equalizer. However, note that the DFE relies on past decisions to remove the post-cursor ISI. Since the decoder decisions are much more reliable than the tentative decisions made by the hard-limiter, this information can be transferred from the channel decoder to the DFE. This facilitates an improved post-cursor cancellation by the DFE. Further, it is sufficient for the channel decoder to provide hard decisions for the DFE. The result is a cleaner soft input for the channel decoder. An interleaver/deinterleaver combination is used to provide some resistance to the bursty nature of errors caused by error propagation. This process can be iterated to improve performance. However, a non-zero delay is associated with the decoding operation. This delay is considered in the turbo DFE design of Figure 6.2.

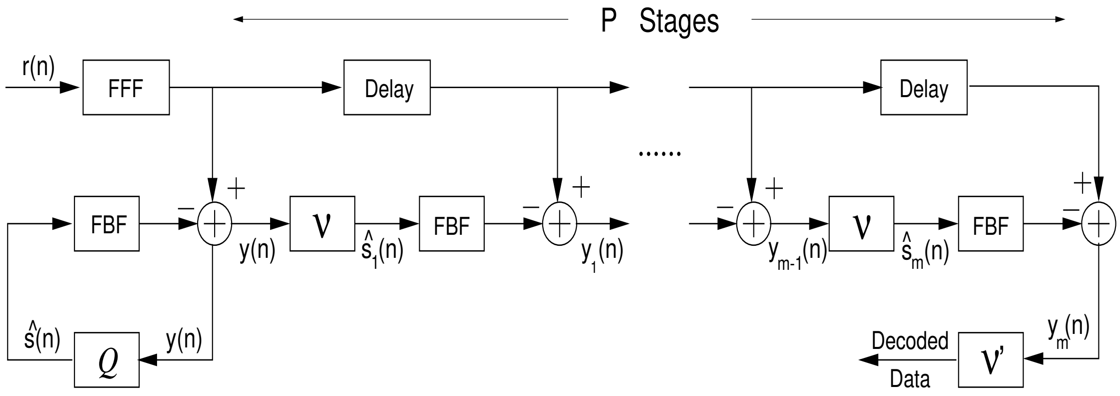

Figure 6.2: Block diagram of a turbo DFE scheme with non-zero internal delay

$r(n)$ are the samples of the received signals with ISI. The precursor ISI is removed with the help of a feed forward filter (FFF). Module $\mathcal{Q}$ outputs tentative decisions (soft/hard), $\hat{s}(n)$, of the transmitted symbols and these are fed back to the feedback filter(FBF) of the DFE to cancel the post-cursor ISI. $y(n)$ are the soft outputs from the DFE block and these are input to the channel decoder. Module $\mathcal{V}$ consists of a deinterleaver followed by the Viterbi Algorithm and an interleaver. The deinterleaver randomizes any bursts of errors and thus provides a marginal resistance to error propagation. Unlike the turbo equalizer, the turbo

DFE does not require a SOVA. Hard decisions are sufficient for the consequent DFE modules. Hence, a hard decision Viterbi algorithm is used as the channel decoder. The reliable decisions of the channel decoder are interleaved before being fed into the next FBF module. Since the decoder introduces a delay, the output of the FFF stage are delayed appropriately before further iterations. With each iteration of the turbo DFE, the reliability of the decisions, $\hat{s}_k(n)$ increases. After a fixed number of iterations, P, module $\mathcal{V}'$ outputs the decoded data. This final module does not have an interleaver in it.

6.3 Simulation Results

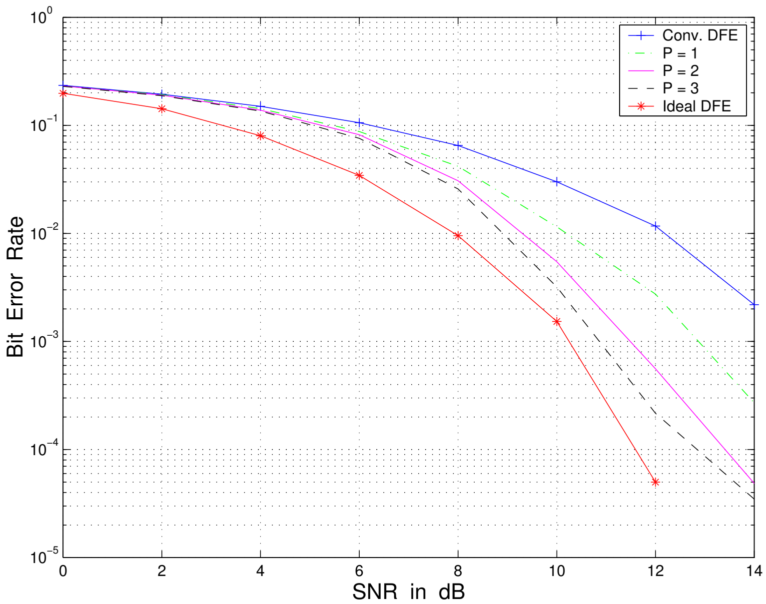

Figure 6.3: BER performance comparison of turbo DFE (with hard decisions) for h = [1 0.05 0.6 -0.27 0.03 0.12 0.22]

A 4-PAM modulation scheme was used for the simulations. A (2,1,3) convolutional code with a generator matrix $[1 + D^2 \quad 1 + D + D^2]$ was used as the error correcting code. The test channels had a combined channel feed-forward filter impulse responses of [1 0.05 0.6 -0.27 0.03 0.12 0.22] and [1 0 0 -0.7]. Perfect knowledge of the combined impulse response was assumed at the receiver. The Viterbi decoder had a decoding delay of atleast 12 symbols. A random interleaver of size 96 was used to randomize the error bursts caused by error propagation.

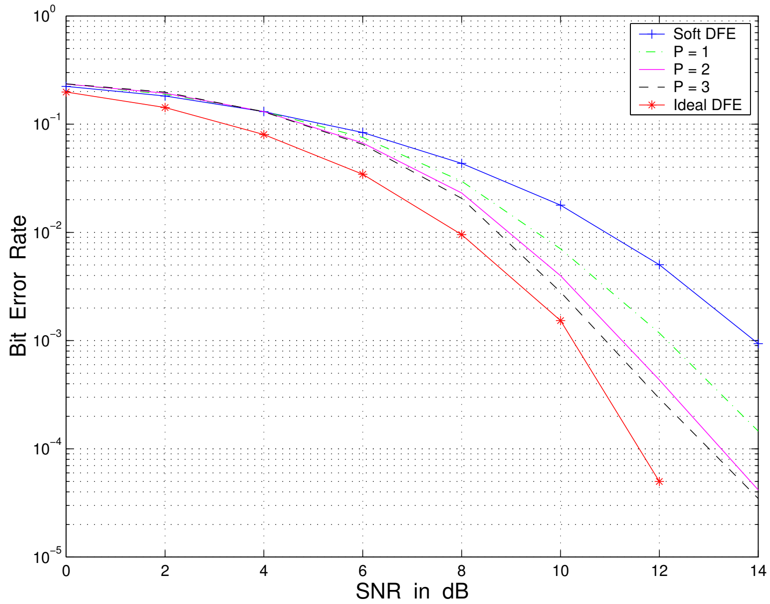

Figure 6.4: BER performance comparison of turbo DFE (with soft decisions) for h = [1 0.05 0.6 -0.27 0.03 0.12 0.22]

Figure 6.3 compares the BER plots of a turbo DFE with that of the ideal DFE in the presence of AWGN for h = [1 0.05 0.6 -0.27 0.03 0.12 0.22]. A conventional DFE was used for this simulation experiment and hence, a hard-limiter was used as module $\mathcal{Q}$ of Figure 6.2 giving rise to tentative hard decisions, $\hat{s}(n)$. At a BER of 10^{-2} the gap between the conventional DFE curve (no iterations) and the ideal DFE curve is about 4.3 dB. With one additional iteration (P = 1) the performance is 2.2 dB away from the ideal DFE bound, with two additional iterations (P = 2) it reaches within 1.5 dB of the ideal DFE bound and for three additional iterations (P = 3) the gap is only 1 dB. These results demonstrate the law of diminishing returns [39] for the addition of more stages of iterations.

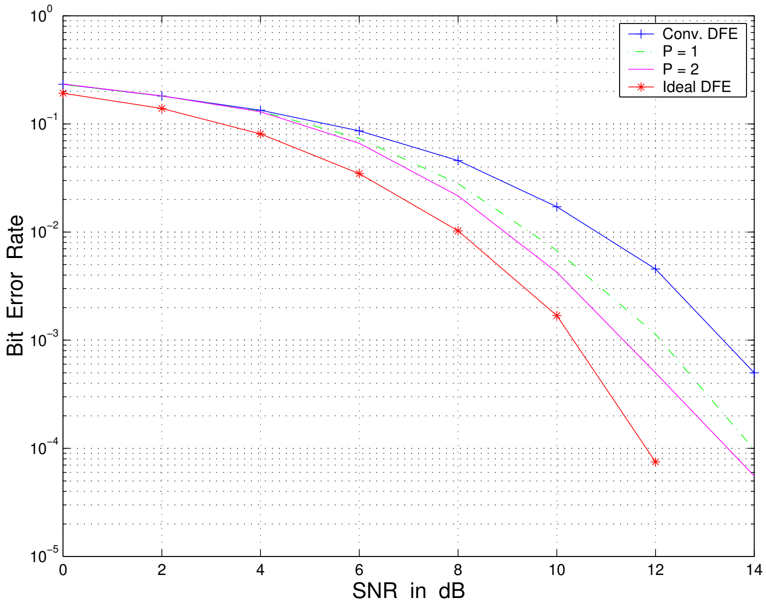

Figure 6.5: BER performance comparison of turbo DFE (with hard decisions) for h = [1 0 0 -0.7]

Figure 6.4 compares the BER plots of a turbo DFE when a soft decision box is used as module $\mathcal{Q}$ of Figure 6.2 giving rise to tentative soft decisions, $\hat{s}(n)$. At a BER of 10^{-2} the gap between the soft DFE curve (no iterations) and the ideal DFE curve is about 3 dB (as opposed to 4.3 dB for the conventional DFE case). With one additional iteration (P = 1) the performance is 1.6 dB away from the ideal DFE bound, with two additional iterations (P = 2) it reaches within 1.1 dB of the ideal DFE bound and for three additional iterations (P = 3) the gap is only

0.9 dB. Though the soft DFE case has better initial gains when compared to the conventional DFE case, the performance difference is marginal for $P \geq 3$. Figures 6.5 and 6.6 illustrates the turbo DFE performance using a conventional DFE and a soft DFE for the other test channel.

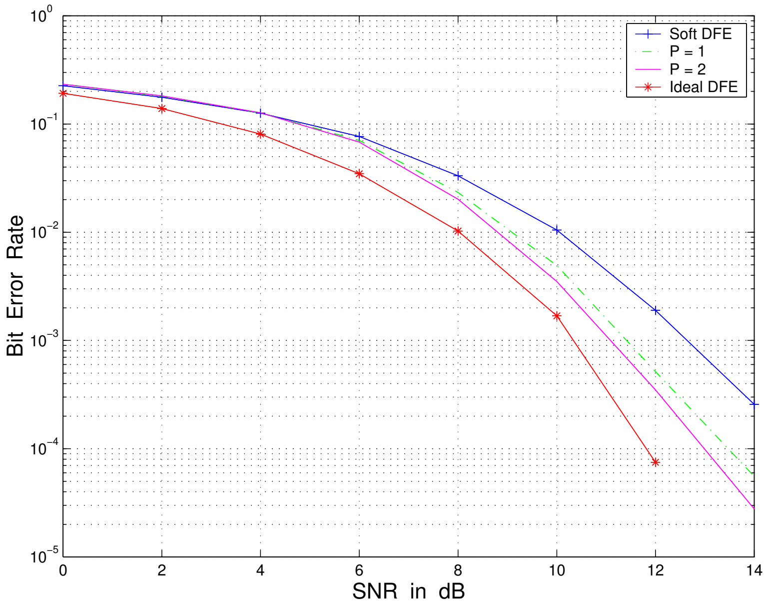

Figure 6.6: BER performance comparison of turbo DFE (with soft decisions) for h = [1 0 0 -0.7]

Chapter 7

Conclusions and Future Directions

7.1 Summary of Results

We have proposed a soft decision feedback approach for mitigating error propagation. The hard-limiter of the conventional DFE was replaced by a soft decision device in an attempt to combat error propagation. The decision device was optimized for the mean squared error (MSE) cost function in the presence of error propagation under various simplifying assumptions on the structure of residual ISI and noise. The optimal decision device was derived for a few specific modulation schemes like QPSK, 8PSK and 16QAM. At high SNR values, the soft decision device was observed to tend towards the hard-limiter of a conventional DFE. Simulation results confirming performance improvements were shown. A bit error rate (BER) optimized decision device was derived for a BPSK modulation scheme in the presence of additive white Gaussian noise. However, the BER optimized non-linearity was observed to be dependent on the value DFE feedback filter coefficient. Further, it had the disadvantage of being in implicit form. A special case of a source with Gaussian distribution was considered. The soft DFE decision device and the DFE taps were jointly optimized (under the MSE criterion) for the Gaussian source.

The soft DFE performance was analyzed by approximating the decision device nonlinearity using a multi-level quantization approximation. The error dynamics of the system was modeled using a Markov chain with $(4K+1)^{L_b}$ states for a $2K+1$-level quantizer and L_b feedback taps for the DFE. The performance measures, namely MSE and BER, were computed from the state occupancy probabilities of the Markov chain. The multi-level quantization approximation was validated by comparing theoretical predictions with simulation results. Two sub-optimal implementations, namely multi-level quantization and piece-wise linear approximation, of the soft DFE non-linearity were explored. Simulation results demonstrating performance improvements of the soft DFE, for a measured microwave channel and a few broadband wireless channel models specified by GSM, were shown.

The need for joint coding and equalization strategies was discussed. The turbo-equalization principle was briefly explained. A turbo DFE strategy based on this idea was proposed to effectively combine DFE with an error correcting decoder. The rationale behind this approach was explained. Simulation results showed improvements over the disjoint equalization and coding scheme. A performance improvement of 2-3 dB was observed for the test channels.

7.2 Future Directions

In this thesis, the soft DFE decision device was optimized under the assumption that the FFF and FBF taps of the DFE are known. The DFE taps were not jointly optimized with the decision device non-linearity. However, a joint optimization of the DFE taps and the decision device, if possible, would be a very useful result.

The Markov-chain based approach for performance analysis of error propagation for the soft DFE is not applicable in real-time in a cost effective manner when the channel has a long memory. Various research papers have derived tight bounds for error propagation by clumping error states of the Markov chain. It would be interesting to investigate if these approaches are equally effective for the soft DFE. Further, due to the feedback structure of the DFE, one should be able to compute the probability distribution function of the error at the input of the decision device in a recursive manner. The feasibility of such an approach needs to be investigated.

The soft DFE motivates research on decision-directed adaptation schemes that use the soft decision errors for the DFE tap coefficient updates instead of the conventional decision-directed adaptation schemes. In fact, the soft DFE shows promise as a candidate for blind adaptation. Since, the decision device non-linearity is a function of the noise variance at the input of the decision box, the soft DFE may behave as a self-learning equalizer.

Appendix A

Derivation of BER Optimized Decision Box

This appendix presents the detailed derivation of the BER optimized non-linearity of the soft DFE decision device for a BPSK modulation scheme. Recall,

$$\begin{aligned}\mathbf{J}(d_n, y_n) = \frac{1}{2}\Bigg[&\Pr(s_n = 1|y_n)\left\{\mathcal{Q}\left(\frac{1-b(1)(1-d_n)}{\sigma_z}\right) + \mathcal{Q}\left(\frac{1+b(1)(1-d_n)}{\sigma_z}\right)\right\}\\ &+\Pr(s_n = -1|y_n)\left\{\mathcal{Q}\left(\frac{1+b(1)(1+d_n)}{\sigma_z}\right) + \mathcal{Q}\left(\frac{1-b(1)(1+d_n)}{\sigma_z}\right)\right\}\Bigg]\end{aligned}$$

To minimize the BER cost function, set

$$\frac{\partial \mathbf{J}(d_n, y_n)}{\partial d_n} = 0 \tag{A.1}$$

Since,

$$\frac{\partial \mathcal{Q}(x)}{\partial x} = \frac{-1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right) \tag{A.2}$$

$$\begin{aligned}\frac{\partial \mathbf{J}(d_n, y_n)}{\partial d_n} = &\frac{b(1)}{2\sqrt{2\pi}\sigma_z}\Pr(s_n = 1|y_n)\left[\exp\left\{\frac{-1}{2\sigma_z^2}\left(1+b(1)(1-d_n)\right)^2\right\}\right.\\ &\left.-\exp\left\{\frac{-1}{2\sigma_z^2}\left(1-b(1)(1-d_n)\right)^2\right\}\right]\\ &+\frac{b(1)}{2\sqrt{2\pi}\sigma_z}\Pr(s_n = -1|y_n)\left[\exp\left\{\frac{-1}{2\sigma_z^2}\left(1-b(1)(1+d_n)\right)^2\right\}\right.\\ &\left.-\exp\left\{\frac{-1}{2\sigma_z^2}\left(1+b(1)(1+d_n)\right)^2\right\}\right]\end{aligned} \tag{A.3}$$

Setting this equal to zero and on simplification,

$$\frac{\Pr(s_n = 1|y_n)}{\Pr(s_n = -1|y_n)} = \frac{\exp\left(\frac{-2b^2(1)d_n}{\sigma_z^2}\right)\sinh\left\{\frac{b(1)(1+d_n)}{\sigma_z^2}\right\}}{\sinh\left\{\frac{b(1)(1-d_n)}{\sigma_z^2}\right\}} \tag{A.4}$$

The input to the decision box can be written as,

$$y(n) = s(n) + noise$$

where noise includes both residual ISI and AWGN. From equation (2.6),

$$noise = z(n) + b(1)\{s(n-1) - d(n-1)\}$$

and under the assumption that *noise* is a zero-mean Gaussian process with variance σ^2 (Assumption 4 of Chapter 2)

$$\begin{aligned}\frac{\Pr(s_n = 1|y_n)}{\Pr(s_n = -1|y_n)} &= \frac{\exp\left\{\frac{-1}{2\sigma^2}(y_n - 1)^2\right\}}{\exp\left\{\frac{-1}{2\sigma^2}(y_n + 1)^2\right\}} \\ &= \exp\left\{\frac{2y_n}{\sigma^2}\right\} \end{aligned} \tag{A.5}$$

Hence, the optimal decision device non-linearity is given by

$$\exp\left(\frac{2y_n}{\sigma^2}\right) = \frac{\exp\left(\frac{-2b^2(1)d_n}{\sigma_z^2}\right)\sinh\left\{\frac{b(1)(1+d_n)}{\sigma_z^2}\right\}}{\sinh\left\{\frac{b(1)(1-d_n)}{\sigma_z^2}\right\}} \tag{A.6}$$

Appendix B

Source Code

Soft_dfe_simulator.m

This function simulates an uncoded system and compares the performance of a soft DFE, conventional DFE, ideal DFE and a linear approximation for a soft DFE.

```
% 'Lf' - length of FFF

% 'Lb' - length of FBF

% 'P' - baud spaced / fractionally-spaced DFE

%%%%%%% Initialisation

rand('seed',0);

ite_num=100000; % Number of iterations

SNR_min=0;SNR_max=16;SNR_step=1; % SNR range

modln=2;  % 0 - BPSK ; 1 - 4PAM ; 2 - QPSK ; 3 - 16QAM

% Choose or generate channel

% C=[1 0.6 -0.55 -0.1];

Lf=8;Lb=8;P=2;

SER_hdfe=[];SER_ideal=[];SER_sdfe=[];SER_sdfe_lin=[];

MSE_hdfe=[];MSE_ideal=[];MSE_sdfe=[];MSE_sdfe_lin=[];
```

```
% Transmitter...
if (modln==0)
 source_seq=pam(ite_num,2,1);
elseif (modln==1)
 source_seq=pam(ite_num,4,1);
elseif (modln==2)
 source_seq=qam(ite_num,4,1);
elseif (modln==3)
 source_seq=qam(ite_num,16,1);
else end;
if (P==1)
 tx_seq=conv(source_seq,C);
elseif (P==2)
 up_seq=zeros(1,2*ite_num);
 up_seq(1:2:2*ite_num)=source_seq;
 tx_seq=conv(up_seq,C);
end;
% End of transmitter part
for snr=SNR_min:SNR_step:SNR_max,
lambda=10^(-snr/10);
% Determine DFE taps
[f_opt,b_opt,delay,mmse]=dfe_tap_estimator(C,Lf,Lb,lambda,P);
if (modln<2)
 rx_seq=tx_seq+randn(1,length(tx_seq))*10^(-snr/20)*sqrt(abs(C*C'));
else
```

```
 rx_seq=tx_seq+(randn(1,length(tx_seq))+j*randn(1,length(tx_seq)))
*10^(-snr/20)*sqrt(abs(C*C')/2);
end;
% Equalization....
eq_seq =conv(rx_seq,f_opt);
if (P==2)
 down_seq=eq_seq(delay:2:length(eq_seq));
elseif (P==1)
 down_seq=eq_seq(delay:length(eq_seq));
end;
%%%%%%%%%%%%%%%%%  Ideal DFE...
ser=0;mse=0;
for k=Lb+1:ite_num,
 est_seq=down_seq(k)-source_seq(k-1:-1:k-Lb)*b_opt;
 mse=mse+(abs(est_seq-source_seq(k)))^2;
 if (modln==0)
  est_sym(k)=pamdet(est_seq,2,1);
 elseif (modln==1)
  est_sym(k)=pamdet(est_seq,4,1);
 elseif (modln==2)
  est_sym(k)=qamdet(est_seq,4,1);
 elseif (modln==3)
  est_sym(k)=qamdet(est_seq,16,1);
 else end;
 err=source_seq(k)-est_sym(k);
```

```
 ser=ser+sign(abs(err));
end;
ser_ideal=ser/(ite_num-Lb);
mse_ideal=mse/(ite_num-Lb);
%%%%%%%%%%%%%%% Conventional DFE
est_sym(1:Lb)=source_seq(1:Lb);
ser=0;mse=0;
for k=Lb+1:ite_num,
 est_seq=down_seq(k)-est_sym(k-1:-1:k-Lb)*b_opt;
 mse=mse+(abs(est_seq-source_seq(k)))^2;
 if (modln==0)
  est_sym(k)=pamdet(est_seq,2,1);
 elseif (modln==1)
  est_sym(k)=pamdet(est_seq,4,1);
 elseif (modln==2)
  est_sym(k)=qamdet(est_seq,4,1);
 elseif (modln==3)
  est_sym(k)=qamdet(est_seq,16,1);
 else end;
 err=source_seq(k)-est_sym(k);
 ser=ser+sign(abs(err));
end;
ser_hdfe=ser/(ite_num-Lb);
mse_hdfe=mse/(ite_num-Lb);
%%%%%%%%%%%%%%%% Soft DFE
```

```
ser=0;mse=0;
%noise_est=10^(-snr/10)*(abs(C*C')*abs(f_opt'*f_opt));
noise_est=abs(mmse);
sdfe_sym(1:Lb)=source_seq(1:Lb);
for k=Lb+1:ite_num,
 est_seq=down_seq(k)-sdfe_sym(k-1:-1:k-Lb)*b_opt;
 mse=mse+(abs(est_seq-source_seq(k)))^2;
 if (modln==0)
  est_sym(k)=pamdet(est_seq,2,1);
  sdfe_sym(k)=tanh(est_seq/noise_est);
 elseif (modln==1)
  est_sym(k)=pamdet(est_seq,4,1);
  alpha=sqrt(1/5);
  sdfe_sym(k)=alpha*(sinh(alpha*est_seq/noise_est)+
   3*exp(-4/5/noise_est)*sinh(3*alpha*est_seq/noise_est))/
   (cosh(alpha*est_seq/noise_est)+exp(-4/5/noise_est)*
   cosh(3*alpha*est_seq/noise_est));
 elseif (modln==2)
  est_sym(k)=qamdet(est_seq,4,1);
  sdfe_sym(k)=(tanh(sqrt(2)*real(est_seq)*10^(snr/10))+
   j*tanh(sqrt(2)*imag(est_seq)*10^(snr/10)))/sqrt(2);
 elseif (modln==3)
  est_sym(k)=qamdet(est_seq,16,1);
  alpha=sqrt(1/10);
  sdfe_sym(k)=alpha*(sinh(2*alpha*real(est_seq)/noise_est)+
```

```
   3*exp(-8/10/noise_est)*sinh(6*alpha*real(est_seq)/noise_est))/
   (cosh(2*alpha*real(est_seq)/noise_est)+exp(-8/10/noise_est)*
   cosh(6*alpha*real(est_seq)/noise_est))+j*alpha*
   (sinh(2*alpha*imag(est_seq)/noise_est)+
   3*exp(-8/10/noise_est)*sinh(6*alpha*imag(est_seq)/noise_est))/
   (cosh(2*alpha*imag(est_seq)/noise_est)+exp(-8/10/noise_est)*
   cosh(6*alpha*imag(est_seq)/noise_est));
 else end;
 err=source_seq(k)-est_sym(k);
 ser=ser+sign(abs(err));
end;
ser_sdfe=ser/(ite_num-Lb);
mse_sdfe=mse/(ite_num-Lb);
%%%%%%%% Linear approximation for SDFE...
ser=0;mse=0;
%noise_est=10^(-snr/10)*(abs(C*C')*abs(f_opt'*f_opt));
noise_est=abs(mmse);
sdfe_lin_sym(1:J)=source_seq(1:J);
for k=J+1:ite_num,
 est_seq=rx_seq(k)/C(1)-sdfe_lin_sym(k-1:-1:k-J)*transpose(weight);
 mse=mse+(abs(est_seq-source_seq(k)))^2;
 if (modln==0)
  est_sym(k)=pamdet(est_seq,2,1);
  sdfe_lin_sym(k)=est_seq/noise_est;
  if (abs(sdfe_lin_sym(k))>1)
```

```
  sdfe_lin_sym(k)=sign(sdfe_lin_sym(k));
 end;
elseif (modln==1)
 est_sym(k)=pamdet(est_seq,4,1);
 est_seq=est_seq*sqrt(5);
   if (abs(est_seq)<1),
  sdfe_lin_sym(k)=est_seq/(5*noise_est);
  if (abs(sdfe_lin_sym(k))>1),
   sdfe_lin_sym(k)=sign(sdfe_lin_sym(k));
  end;
 else
  sdfe_lin_sym(k)=(est_seq-2*(sign(est_seq)))/(5*noise_est)
   +(sign(est_seq))*2;
  if (abs(est_seq-2*(sign(est_seq)))/(5*noise_est) > 1)
   sdfe_lin_sym(k)=sign(est_seq-2*(sign(est_seq)))+2*sign(est_seq);
  end;
 end;
 sdfe_lin_sym(k)=sdfe_lin_sym(k)/sqrt(5);
 elseif (modln==2)
  est_sym(k)=qamdet(est_seq,4,1);
  temp_real=real(est_seq)/noise_est;
  if (abs(temp_real)>1/sqrt(2))
   temp_real=sign(temp_real)/sqrt(2);
  end;
  temp_imag=imag(est_seq)/noise_est;
```

```
 if (abs(temp_imag)>1/sqrt(2))
  temp_imag=sign(temp_imag)/sqrt(2);
 end;
 sdfe_lin_sym(k)=temp_real+j*temp_imag;
elseif (modln==3)
 est_sym(k)=qamdet(est_seq,16,1);
 est_seq=est_seq*sqrt(10);
  if (abs(real(est_seq))<1),
   temp_real=real(est_seq)/(5*noise_est);
   if (abs(temp_real)>1),
    temp_real=sign(temp_real);
   end;
  else
   temp_real=(real(est_seq)-2*(sign(real(est_seq))))/
    (5*noise_est)+(sign(real(est_seq)))*2;
   if (abs(real(est_seq)-2*(sign(real(est_seq))))/(5*noise_est)>1)
    temp_real=sign(real(est_seq)-2*(sign(real(est_seq))))+
     2*sign(real(est_seq));
   end;
  end;
  if (abs(imag(est_seq))<1),
   temp_imag=imag(est_seq)/(5*noise_est);
   if (abs(temp_imag)>1),
    temp_imag=sign(temp_imag);
   end;
```

```
    else
     temp_imag=(imag(est_seq)-2*(sign(imag(est_seq))))/(5*noise_est)
      +(sign(imag(est_seq)))*2;
     if (abs(imag(est_seq)-2*(sign(imag(est_seq))))/(5*noise_est)>1)
      temp_imag=sign(imag(est_seq)-2*(sign(imag(est_seq))))+
       2*sign(imag(est_seq));
     end;
    end;
   sdfe_lin_sym(k)=(temp_real+j*temp_imag)/sqrt(10);
  else end;
  err=source_seq(k)-est_sym(k);
  ser=ser+sign(abs(err));
end;
ser_sdfe_lin=ser/(ite_num-J);
mse_sdfe_lin=mse/(ite_num-J);
SER_hdfe=[SER_hdfe ser_hdfe];
SER_sdfe=[SER_sdfe ser_sdfe];
SER_ideal=[SER_ideal ser_ideal];
SER_sdfe_lin=[SER_sdfe_lin ser_sdfe_lin];
MSE_hdfe=[MSE_hdfe mse_hdfe];
MSE_sdfe=[MSE_sdfe mse_sdfe];
MSE_ideal=[MSE_ideal mse_ideal];
MSE_sdfe_lin=[MSE_sdfe_lin mse_sdfe_lin];
end;
snr=SNR_min:SNR_step:SNR_max;
```

dfe_tap_estimator.m

This function determines the mmse optimal dfe taps (without including error propagation effects).

```
function [f_opt,b_opt,delay,mmse]=dfe_tap_estimator
 (channel,Lf,Lb,lambda,P);
% 'lambda' is (Noise variance)/(Signal Variance)
% 'P' is the oversampling factor
[h]=channel; Lh=length(h);
cm = transpose(convmtx(h,Lf)); % Find convolution matrix
if (P==2) % T/2 spaced case
  cm1=cm(1:2:Lf+Lh-1,:);
elseif (P==1) cm1=cm;
else end;
  [delta1,junk]=size(cm1);
  I_delta1=eye(delta1);
  f1=zeros(Lf,delta1);
  b1=zeros(Lb,delta1);
  for delay1=1:delta1, % Delay optimization
    I_delta=I_delta1;
    I_delta(delay1+1:min(delta1,delay1+Lb),delay1+1:
     min(delta1,delay1+Lb))=zeros(min(Lb,delta1-delay1));
    m1=inv(cm1'*I_delta*cm1+lambda*eye(Lf))*cm1'*I_delta ;
    temp=I_delta*cm1*m1;
    mse1(delay1)=1-temp(delay1,delay1); % MSE calculation
    f1(:,delay1)=m1(:,delay1); % Feedforward equalizer taps
```

```
  temp=(I_delta1-I_delta)*cm1*f1(:,delay1);
  b1(1:min(Lb,delta1-delay1),delay1)=
   temp(delay1+1:min(delta1,delay1+Lb));
end;
[mmse,index]=min(mse1);  % Delay optimization of mse
  f_opt=f1(:,index); % FFF taps
  b_opt=b1(:,index); % FBF taps
delay=index;
```

Bibliography

[1] R.W. Lucky, "Automatic equalization for digital communications," *Bell System Technical Journal*, Vol.44, pp.547-588, April 1965.

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

[3] P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Transactions on Information Theory*, Vol.17, pp.56-64, January 1971.

[4] M. Tomlinson, "New automatic equalizers employing modulo arithmetic," *Electronics Letters*, Vol.7, No.5-6, pp.138-139, March 1971.

[5] G.D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of inter symbol interference," *IEEE Transactions on Information Theory*, Vol.18, No.3, pp.989-995, May 1972.

[6] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, Vol.29, pp.774-780, August 1972.

[7] G.D. Forney, Jr., "The Viterbi Algorithm," *Proceeding of the IEEE*, Vol.61, No.3, pp.268-278, March 1973.

[8] J. Salz, "Optimum mean-square decision feedback equalization," *Bell System Technical Journal*, Vol.52, No.8, pp.1341-1373, October 1973.

[9] D. Messerschmitt, "A Geometric theory of intersymbol interference. Part I: Zero-forcing and decision feedback equalization," *Bell System Technical Journal*, Vol.52, No.10, pp.1483-1519, November 1973.

[10] D.L. Duttweiler, J.E. Mazo and D.G. Messerschmitt, " An upper bound on the error probability in decision feedback equalization," *IEEE Transactions on Information Theory*, Vol.20, No.4, pp.490-497, July 1974.

[11] W.U. Lee and F.S. Hill, Jr., "A maximum-likelihood sequence estimator with decision-feedback equalization," *IEEE Transactions on Communications*, Vol.25, No.9, pp.971-979, September 1977.

[12] C.A. Belfiore and J.H. Park Jr., "Decision feedback equalization," *Proceedings of the IEEE*, Vol.67, No.8, pp.1143-1156, August 1979.

[13] S.U.H. Qureshi, "Adaptive Equalization," *Proceedings of the IEEE*, Vol.73, No.9, pp.1349-1387, September 1985.

[14] J.J. O'Reilly and A.M. de Oliveira Duarte, "Error propagation in decision feedback equalizers," *Proceedings of the IEEE*, Vol.132, No.7, pp.561-566 December 1985.

[15] A.M. de Oliveira Duarte and J.J. O'Reilly, "Simplified technique for bounding error statistics for DFB receivers," *Proceedings of the IEEE*, Vol.132, No.7, pp.567-575, December 1985.

[16] R.A. Kennedy and B.D.O. Anderson, "Recovery times of decision feedback equalizers on noiseless channels," *IEEE Transactions on Communications*, Vol.35, No.10, pp.1012-1021, October 1987.

[17] R.A. Kennedy, B.D.O.Anderson and R.R. Bitmead "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Transactions on Communications*, Vol.35, No.10, pp.1022-1028, October 1987.

[18] M.V. Eyuboğlu and S.U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Transactions on Communications*, Vol. 36, pp.13-20, January 1988.

[19] M.V. Eyuboğlu, "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interleaving," *IEEE Transactions on Communications*, Vol. 36, pp.401-409, April 1988.

[20] R.A. Kennedy, B.D.O. Anderson and R.R.Bitmead, "Channels leading to rapid error recovery for decision feedback equalizers," *IEEE Transactions on Communications*, Vol.37, No.11, pp.1126-1135, November 1989.

[21] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," *Proc. IEEE GLOBECOM '89*, Dallas, Texas, November 1989.

[22] K. Zhou, J.G. Proakis and F. Ling, "Decision feedback equalization of time-dispersive channels with coded modulation," *IEEE Transactions on Communications*, Vol.38, No.1, pp.18-24, January 1990.

[23] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboğlu and G.D. Forney, Jr., "MMSE Decision-Feedback Equalizers and Coding-Part I: Equalization Results," *IEEE Transactions on Communications*, Vol. 43, No.10, pp.2582-2594, October 1990.

[24] S.A. Altekar and N.C. Beaulieu, "Upper bounds to the error probability of decision feedback equalization," *IEEE Transactions on Information Theory*, Vol.39, No.1, pp.145-156, January 1993.

[25] N.C. Beaulieu, "Bounds on Recovery Times of Decision Feedback Equalizers," *IEEE Transactions on Communications*, Vol.42, No.10, pp.2786-2794, October 1994.

[26] J.E. Smee, "New methods for evaluating equalizer error rate performance," *IEEE Vehicular Technology Conference*, Vol.1, pp.87-91, July 1995.

[27] N. Al-Dhahir and J.M. Cioffi, "MMSE Decision-Feedback Equalizers: Finite-Length Results," *IEEE Transactions on Information Theory*, Vol.41, No.4, pp.961-975, July 1995.

[28] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier and A. Glavieux "Iterative correction of intersymbol interference: turbo equalization," *European Transactions on Telecommunications*, Vol.6, pp.507-511, September 1995.

[29] S.H. Muller, W.H. Gerstacker and J.B. Huber, "Reduced-state soft-output trellis-equalization incorporating soft-feedback," *IEEE GLOBECOM '96*, pp.95-100, London, UK, November 1996.

[30] O.E. Agazzi, "On the use of tentative decisions to cancel intersymbol interference and nonlinear distortion (with application to magnetic recording channels)," *IEEE Transactions on Information Theory*, Vol.43, No.2, pp.394-408, March 1997.

[31] D. Yellin, A. Vardy and O. Amrani, "Joint Equalization and Coding for Intersymbol Interference Channels," *IEEE Transactions on Information Theory*, Vol.43, No.2, pp.409-425, March 1997.

[32] K. Dogancay and R.A. Kennedy, "Blind detection of equalization errors in communication systems," *IEEE Transactions on Information Theory*, Vol.43, No.2, pp.469-482, March 1997.

[33] R.A. Casas and C.R. Johnson, Jr.,"On the blind adaptation of an FSE+DFE combination," *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Paris, France, April 1997.

[34] R.A. Kennedy, B.D.O. Anderson and R.R. Bitmead, "Stochastic Analysis of Non-Adaptive Decision Feedback Equalizers," *Proceedings of ISSPA87*, pp.842-847, Brisbane, Australia, August 1997.

[35] J.E. Smee and N.C. Beaulieu, "Error-Rate Evaluation of Linear Equalization and Decision Feedback Equalization with Error Propagation," *IEEE Transactions on Communications*, Vol.46, No.5, pp.656-665, May 1998.

[36] S. Ariyavisitakul and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE Vehicular Technology Conference*, Vol.3, pp.2256-2261, May 1998.

[37] S. Ariyavisitakul and G.M. Durant, "A broadband wireless packet technique based on coding, diversity and equalization," *IEEE Communications Magazine*, Vol.36, No.7, pp.110-115, July 1998.

[38] A. Fertner, "Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation," *IEEE Transactions on Signal Processing*, Vol.46, No.7, pp.1872-1877, July 1998.

[39] P.A. Samuelson, "Economics," *McGraw Hill*, 1970.

[40] J.G. Proakis, "Digital Communications," *McGraw Hill*, 1989.

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