Decision Device Optimization for Soft Decision Feedback Equalization

Jaiganesh Balakrishnan†
Cornell University
Ithaca, NY 14850

Harish Viswanathan
Lucent Technologies Bell Labs
Holmdel, NJ

C. Richard Johnson, Jr.,
Cornell University
Ithaca, NY 14850

Abstract — **This paper considers the use of soft decisions to mitigate error propagation in a decision feedback equalizer (DFE). The decision device for the soft decision feedback equalizer is optimized by minimizing the mean-squared error (MSE) cost function defined at the input of the decision device. Furthermore, a bit error rate (BER) optimized soft decision device is derived for a BPSK modulation scheme. Simulation results comparing the performance of the MSE/BER optimized soft DFE and the conventional DFE are presented.**

I. Introduction

Future generation wireless communication technology promises an increase in data rate capabilities of present radio links. As we go for higher rates the channel becomes increasingly frequency selective. The resulting signal distortion caused by the overlap of successive symbols in the received signal is called inter-symbol interference (ISI), which, if left uncompensated, causes high error rates. One suboptimal scheme that is widely used in practice, due to its simple implementation, is the linear equalizer. But, in suppressing the inter-symbol interference, the linear equalizer inevitably enhances the channel noise. This basic limitation of a linear equalizer's ability 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 (DFE).

The idea of using previous decisions to cope with the ISI problem was first introduced in 1967 by Austin [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 [2]. The DFE design has always been carried out assuming that the past decisions are error-free, thus simplifying the mathematics involved. However, when an error is made by the receiver, the output of the feedback filter (FBF) of the DFE 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.

Although the phenomenon of error propagation has been studied extensively, for example [5, 7, 8], very little work has been done on mitigating error propagation in decision feedback equalizers. Techniques for detecting decision errors in equalization schemes including DFEs were proposed in [10]. Several methods, based on transmitter precoding [3, 4], were proposed in the early 1970s to remedy the error propagation problem in the DFE. These methods 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. A scheme which attempts to prevent primary errors and hence error propagation was introduced in [12]. The performance of this algorithm critically depends on the value of the first post-cursor and the accompanying noise level. The use of a combination of soft decisions and delayed tentative decisions to cancel ISI for a joint convolutional coding and DFE scheme was proposed in [11]. The soft decision device is obtained by using a simplified maximum aposteriori probability (MAP) algorithm.

In this paper the idea of using soft decisions to combat error propagation is considered. Section II describes the system model and the notation used in this paper. The decision device of the soft DFE is optimized by minimizing the mean-squared error (MSE) cost function defined at the input of the decision device, under various simplifying assumptions in Section III. In Section IV, the soft decision device under the minimum bit error rate (BER) criterion is derived for a BPSK modulation scheme. Simulation results comparing the performance of the MSE/BER optimized soft DFE and the conventional DFE are presented in Section V.

II. System Model

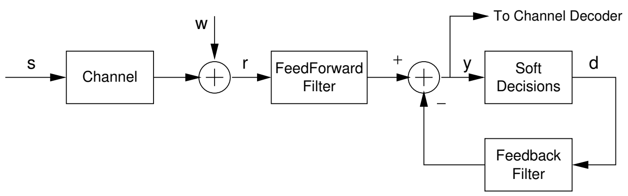

Figure 1: Digital baseband equivalent of the system

†This work was supported by Lucent Technologies and NSF grant ECS-9528363

Figure 1 depicts the digital baseband equivalent of the system. 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 and is modeled as a linear FIR filter. The transmitted signal is furthermore corrupted by additive white Gaussian noise (AWGN) at the receiver.

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

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

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

$$y(n) = \mathbf{s}^T\mathbf{h} + \mathbf{w}^T\mathbf{f} - \mathbf{d}^T\mathbf{b} \qquad (2)$$

where $\mathbf{w} = [w(n)\ w(n-1)\ \dots\ w(n-L_f+1)]^T$ is the noise vector, $\mathbf{d} = [d(n-1)\ d(n-2)\ \dots\ d(n-L_b)]^T$ is the vector containing the past soft decisions and $\mathbf{s} = [s(n+\delta)\ \dots\ s(n+\delta-L_h+1)]^T$ is the transmitted symbol vector. The feedback filter subtracts the contribution of the L_b taps of $\mathbf{h}$ that follow the cursor.

Assumption 1: The "optimal" feedforward and feedback filter taps are known at the receiver. It is also assumed that $\mathbf{b} = [h(\delta+1)\ h(\delta+2)\ \dots\ h(\delta+L_b)]^T$ 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 the goal here is to optimize only the decision device, the above assumption is made. Various techniques for estimating the optimal DFE taps have been discussed in the literature, for example [2, 6, 9]. Furthermore, 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, usually, the soft inputs $y(n)$ are fed into an error correcting decoder for decoding. The tentative soft decisions $d(n)$ are used as inputs to the feedback filter to cancel ISI.

III. MSE Optimization of the Decision Device

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 application of interest.

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 the 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 is formulated as,

$$\Psi_{opt} = \arg\min_{\Psi} \mathbf{E}\left[|y(n) - s(n)|^2\right] \qquad (3)$$

where $\mathbf{E}[(|y(n) - s(n)|)^2]$ is the MSE at the input of the decision device. For a given soft input $y(n)$, the soft decision $d(n)$ is

$$d(n) = \Psi\{y(n)\} \qquad (4)$$

Let the input to the decision device at the n $^{\text{th}}$ time instant be $y(n)$ and the corresponding soft decision be $d(n)$. 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+$\text{L}_\text{b}^{\text{th}}$ time instant. The optimization problem is formulated as determining the optimal soft decision $d(n)$ that would minimize the expected value of squared error at the decision device input during the n+1 $^{\text{th}}$ 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 $^{\text{th}}$ time instant is ,

$$\mathbf{J}(d_n, y_n) = \sum_{S\in\mathcal{A}} \Pr(s_n = S|y_n) \cdot \mathbf{E}\left[(|y_{n+1} - s_{n+1}|)^2 | s_n = S, y_n\right] \qquad (5)$$

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

$$y_{n+1} = s_{n+1} + b(1)[s_n - d_n] + z_{n+1} \qquad (6)$$

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

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

$$z_n = \sum_{i=2}^{L_b} b(i)[s_{n-i} - d_{n-i}] + \sum_{i=0}^{L_f} f(i) w_{n-i} + \sum_{i=0}^{i=\delta-1} h(i) s_{n+\delta-i} + \sum_{i=\delta+L_b+1}^{L_h} h(i) s_{n+\delta-i} \tag{7}$$

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. Furthermore, if the decision device were to be symmetric about the origin, the modeled post-cursor ISI term will also be zero mean. As shall be seen later, the decision device is symmetric about the origin for a symmetric constellation.

Assumption 2: $z(n+1)$ 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.

From equation (6),

$$\begin{aligned}
&\mathbf{E}\left[(|y_{n+1} - s_{n+1}|)^2 | s_n = S, y_n\right] \\
&= \mathbf{E}\left[\{|b(1)\{s_n - d_n\} + z_{n+1}|\}^2 | s_n = S, y_n\right] \\
&= |b(1)\{S - d_n\}|^2 + \mathbf{E}\left[\{|z_{n+1}|\}^2 | s_n = S, y_n\right] \\
&\quad + \mathbf{E}\left[b(1)\{S - d_n\} z^*_{n+1} | y_n\right] \\
&\quad + \mathbf{E}\left[b^*(1)\{S^* - d^*_n\} z_{n+1} | y_n\right]
\end{aligned} \tag{8}$$

Due to Assumption 2, the last two terms go to zero and

$$\begin{aligned}
&\mathbf{E}\left[(|y_{n+1} - s_{n+1}|)^2 | s_n = S\right] \\
&\quad = \{|b(1)\{S - d(n)\}|\}^2 \\
&\quad\quad + \mathbf{E}\left[\{|z(n+1)|\}^2 | s_n = S, y_n\right]
\end{aligned} \tag{9}$$

Note that the last term is independent of d_n. Hence equation (5) reduces to

$$\mathbf{J}(d_n, y_n) = \sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n) \cdot \Big[|b(1)\{S - d_n\}|^2 + \mathbf{E}\left[|z_{n+1}|^2 | s_n = S, y_n\right] \Big] \tag{10}$$

Minimizing the MSE cost function,

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

$$\Rightarrow \quad \sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n)[S - d_n] = 0 \tag{11}$$

$$\Rightarrow \quad d_n = \sum_{S \in \mathcal{A}} \Pr(s_n = S|y_n) S \tag{12}$$

From equation (12) the optimal decision device is given by,

$$\Psi_{opt}(y_n) = \mathbf{E}[s_n | y_n] \tag{13}$$

In [11] Ariyavisitakul and Li obtain the same expression, i.e., equation (13), for the soft decision device, but by approximating the optimum MAP algorithm that minimizes the mean-squared error between s(n) and d(n).

The input to the soft decision device can be written as

$$y(n) = s(n) + \eta(n) \tag{14}$$

where noise includes both residual ISI and AWGN. From equation (6),

$$\eta(n) = z(n) + b(1)\{s(n-1) - d(n-1)\} \tag{15}$$

Assumption 3: $\eta(\cdot)$ is a zero mean random process and has a complex Gaussian distribution
From Claim 1 we can conclude that noise is zero mean. Furthermore, if the residual ISI has a reasonably large number of small terms then central limit theorem can be invoked and a Gaussian approximation can be assumed.

Let $\eta(\cdot)$ have a total variance of σ^2. Under the above mentioned assumption, the decision device input/output relationship for any linear two-dimensional constellation with equi-probable symbols is,

$$\Psi_{opt}(y_n) = \frac{\sum_{S \in \mathcal{A}} S \exp\left[\frac{-S^* S + y_n^* S + S^* y_n}{\sigma^2}\right]}{\sum_{S \in \mathcal{A}} \exp\left[\frac{-S^* S + y_n^* S + S^* y_n}{\sigma^2}\right]} \cdot \tag{16}$$

In particular, for a QPSK constellation, $\mathcal{A} = \left\{\pm\frac{1}{\sqrt{2}} \pm \frac{j}{\sqrt{2}}\right\}$ the decision device nonlinearity is,

$$\Psi_{opt}(y_n) = \frac{1}{\sqrt{2}} \tanh\left(\frac{\sqrt{2}\mathrm{Re}\{y(n)\}}{\sigma^2}\right) + \frac{j}{\sqrt{2}} \tanh\left(\frac{\sqrt{2}\mathrm{Im}\{y(n)\}}{\sigma^2}\right) \tag{17}$$

Similarly, the decision device input/output relationship for a linear one-dimensional constellation with equi-probable symbols is given by,

$$\Psi_{opt}(y_n) = \frac{\sum_{S \in \mathcal{A}} S \exp\left[\frac{-S^2 + 2Sy_n}{2\sigma^2}\right]}{\sum_{S \in \mathcal{A}} \exp\left[\frac{-S^2 + 2Sy_n}{2\sigma^2}\right]} \cdot \tag{18}$$

IV. BER Optimization of the Decision Device

In this section the optimization of the soft decision device to minimize BER rather than MSE is considered. Such a soft

DFE would be of interest in an uncoded scenario, or in a coded scenario where only a hard input channel decoder is used. In deriving the BER optimized decision device the source symbols are assumed to be from a BPSK constellation. The BER optimization problem is formulated as,

$$\Psi_{opt} = \arg\min_{\Psi} \Pr[\text{sign}\{y(n)\} \neq s(n)] \tag{19}$$

where sign$\{y(n)\}$ is the bit decision[2]. The BER cost function is,

$$\mathbf{J}(d_n, y_n) = \sum_{S\in\{\pm1\}} \Pr(s_n = S|y_n)\mathcal{P}(d_n, s_n = S) \tag{20}$$

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

$$\begin{aligned}\mathcal{P}(d_n, s_n = S) = &\sum_{S'\in\{\pm1\}} \Pr(s_{n+1} = S') \\ &\cdot \Pr[\text{sign}\{y_{n+1}\} \neq S'|s_n = S, s_{n+1} = S']\end{aligned} \tag{21}$$

Recall equation (6)

$$y_{n+1} = s_{n+1} + b(1)[s_n - d_n] + z_{n+1}$$

Assuming $z(\cdot)$ to be a zero-mean complex Gaussian random process (similar to **Assumption 3**) with variance σ_z^2

$$\begin{aligned}&\mathcal{P}(d_n, s_n = S) \\ &\quad= \Pr(s_{n+1} = -1)\mathcal{Q}\left(\frac{1 - b(1)(s_n - d_n)}{\sigma_z}\right) \\ &\quad\quad+ \Pr(s_{n+1} = 1)\mathcal{Q}\left(\frac{1 + b(1)(s_n - d_n)}{\sigma_z}\right)\end{aligned} \tag{22}$$

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

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp(\frac{-Z^2}{2})dz \tag{23}$$

If the transmitted symbols are equi-probable

$$\begin{aligned}\mathcal{P}(d_n, s_n = S) = \frac{1}{2}\Bigg\{ &\mathcal{Q}\left(\frac{1 - b(1)(S - d_n)}{\sigma_z}\right) \\ &+\mathcal{Q}\left(\frac{1 + b(1)(S - d_n)}{\sigma_z}\right)\Bigg\}\end{aligned} \tag{24}$$

Substituting equation (24) in equation (20) and minimizing the BER cost function, i.e.,

$$\frac{\partial \mathbf{J}(d_n)}{\partial d_n} = 0 \tag{25}$$

[2]These are the symbol hard-decisions and are different from the tentative soft decisions used by the feedback filter to cancel the post-cursor ISI.

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

$$\exp\left(\frac{2y_n}{\sigma^2}\right) = \frac{\exp\left(\frac{-2b^2(1)d_n}{\sigma_z^2}\right)\sinh\left\{\frac{b(1)(1+d_n)}{\sigma_z^2}\right\}}{\sinh\left\{\frac{b(1)(1-d_n)}{\sigma_z^2}\right\}} \tag{26}$$

In reality, it would not be possible to independently compute σ_z^2. Hence, we assume that $\sigma_z^2 \approx \sigma^2$. The non-linearity of the BER optimized decision device depends on the value of the feedback filter's coefficient $b(1)$. This is due to the fact that we took into account only the effect of the first residual post-cursor ISI term and clumped all other residual post-cursor ISI terms with the noise component. Hence, for best results, the decision device has to be optimized by considering the residual post-cursor ISI term contributed by the dominant feedback filter coefficient. This dependence on the feedback-filter coefficient is a major disadvantage in adaptive scenarios, as the decision device needs to be periodically updated. It is furthermore complicated by the fact that the decision device input/output relation is available only in implicit form. However, the BER optimized soft decision device can be easily implemented by using a piece-wise linear approximation for the decision device non-linearity.

The piece-wise linear approximation of the BER optimized soft decision device input/output relationship is given by,

$$d(n) = \begin{cases} \frac{y(n)}{\gamma} & \text{if } \frac{|y(n)|}{\gamma} \leq 1 \\ \text{sign}[y(n)] & \text{if } \frac{|y(n)|}{\gamma} > 1 \end{cases} \tag{27}$$

where

$$\gamma = b(1)\coth\left\{\frac{b(1)}{\sigma^2}\right\} - b^2(1) \tag{28}$$

and γ is the gradient $\left(\frac{\partial d_n}{\partial y_n}\right)$ of equation (26) evaluated at the origin.

V. Simulation Results

Figure 2 compares the simulation based MSE performance curves of the MSE optimized soft DFE with a conventional DFE and an ideal DFE for a 4-PAM modulation scheme. In an ideal DFE correct decisions are fed back to the feedback filter and hence error propagation does not occur. A combined channel feedforward filter impulse response of

$$\text{h} = [1\ 0.05\ 0.6\ -0.27\ 0.03\ 0.12\ 0.22]$$

was used for this simulation. The variance σ^2 of the noise process $\eta(\cdot)$ of equation (15) was estimated at the receiver by using the update rule,

$$\sigma^2(n+1) = \lambda\sigma^2(n) + (1-\lambda)\{y(n) - \hat{s}(n)\}^2 \tag{29}$$

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

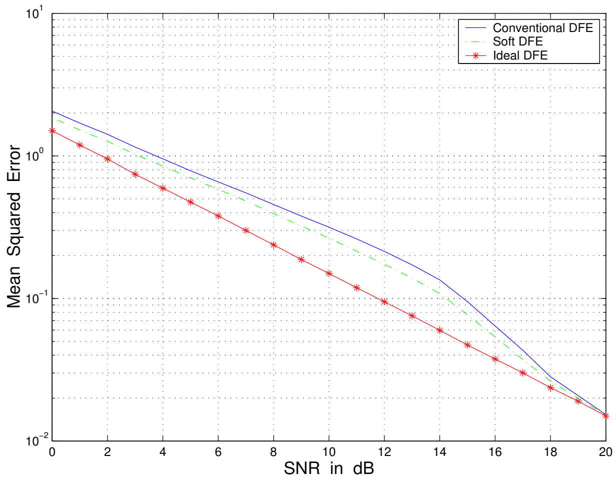

Figure 2: Performance of an MSE optimized soft DFE

the data. σ^2 was initialized with the variance of AWGN. In practical scenarios this initial value can be estimated during the training period. Performance results indicate a reasonable error propagation mitigation for the soft DFE. At higher SNRs, primary errors (and hence the error bursts) occur less frequently and the MSE of the conventional DFE and the soft DFE are almost identical to that of the ideal DFE.

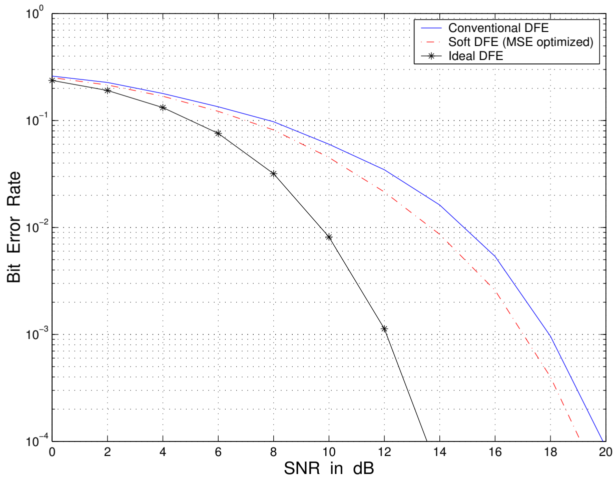

Figure 3: BER performance of an MSE optimized soft DFE for a coded system

Figure 3 compares the BER performance of the MSE optimized soft DFE and the conventional DFE. The same system parameters as that of Figure 2 were assumed. A rate (1/2) convolutional code with a constraint length of 3 was used for the error correcting code. The generator polynomials of the convolutional 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. For example, at a BER of 10^{-2} the error propagation gap is about 5.2 dB and the soft DFE provides an improvement of about 1.3 dB over the conventional DFE. These simulation results demonstrate the advantage of using the soft DFE over the conventional DFE.

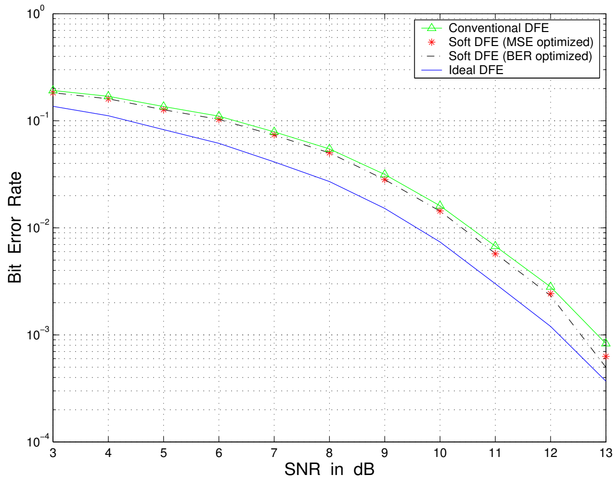

Figure 4: Performance of a BER optimized soft DFE for an uncoded system

Figure 4 compares the BER performance of a BER optimized soft DFE with a conventional DFE, ideal DFE and an MSE optimized soft DFE. The transmitted source symbols were assumed to be from a BPSK constellation. A combined channel feedforward filter impulse response of h = [1 0.6 -0.55 -0.1] was chosen. Although, the BER optimized soft DFE performs better than the conventional DFE, its performance is only marginally better than that of the MSE optimized soft DFE.

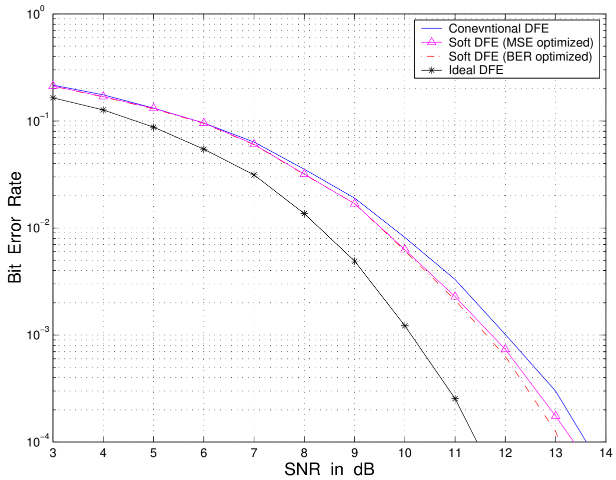

Figure 5: Performance of a BER optimized soft DFE for a coded system

Figure 5 compares the BER performance of the BER optimized soft DFE with a conventional DFE and an MSE optimized soft DFE. The same system parameters as that of Figure 4 were used. A rate (1/2) convolutional code with a constraint length of 3 was assumed as the error correcting code. The generator polynomials of the convolutional code were $g_1(D) = D^2 + 1$ and $g_2(D) = D^2 + D + 1$. A hard input Viterbi decoder with a decoding delay of atleast 12 symbols was used. At a BER of 10^{-3} the error propagation gap is about 2 dB and the BER optimized soft DFE provides an improvement of about 0.5 dB over the conventional DFE. Note that the MSE optimized soft DFE performance is only marginally worse than that of the BER optimized soft DFE. This suggests that an MSE optimized soft DFE is near optimal even under a minimum BER criterion.

VI. Conclusions

The idea of using soft decisions to mitigate error propagation in a DFE was considered in this paper. The MSE optimized soft DFE decision device was derived under various simplifying assumptions. It was found that the MSE optimized soft decision device was the same as the soft decision device nonlinearity derived in [11] by approximating the MAP algorithm that minimizes the mean-squared error between the transmitted symbol and the soft decision. Furthermore, the BER optimized soft DFE decision device was derived for a BPSK modulation scheme. Simulation results demonstrating performance improvements of the MSE optimized soft DFE and the BER optimzed soft DFE over a conventional DFE were presented.

References

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

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

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

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

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.