

# Mitigating Error Propagation Effects in a Decision Feedback Equalizer

Michael Reuter, *Member, IEEE*, Jeffery C. Allen, *Member, IEEE*, James R. Zeidler, *Fellow, IEEE*, and Richard C. North, *Member, IEEE*

**Abstract**—We present an approximate analysis approach to the computation of probability of error and mean burst error length for a decision feedback equalizer (DFE) that takes into account feedback of decision errors. The method uses a reduced-state Markov model of the feedback process and is applicable to linear modulation formats. We use this technique to analyze a DFE design that mitigates the effects of feedback error by incorporating a soft decision device into the feedback path and a norm constraint on the feedback filter weights. We apply the DFE design and analysis approach to a dispersive multipath propagation environment.

**Index Terms**—Decision feedback equalizers, error analysis, error propagation, Markov processes.

## I. INTRODUCTION

THE DECISION feedback equalizer (DFE) is an important component in many digital communication receivers and is used to suppress intersymbol interference (ISI) caused by dispersive propagation channels [1], [2], as well as reject in-band interference [3], [4]. The DFE incorporates a feedforward filter that operates on the received signal to suppress precursor ISI, with a feedback filter that operates on previously detected channel symbols to suppress postcursor ISI. The DFE generally outperforms the traditional linear equalizer, particularly if the channel has deep spectral nulls in its response [1].

However, degradation in DFE performance occurs when incorrectly detected symbols are fed through the feedback filter. Then instead of mitigating ISI from the cursor sample, the DFE enhances ISI. Error propagation may result that causes bursts of decision errors and a corresponding increase in the average probability of bit and symbol error. Moreover, the bursty nature of DFE errors has implications for error correction coding and interleaver depth that may be incorporated into the receiver design [5], [6].

Error propagation generally does not severely affect DFE performance if the channel delay spread is on the order of a symbol duration or less [1], [7]. However, as bandwidths increase to accommodate high data rates, the associated increase in relative delay spread may result in error propagation for a given channel.

Paper approved by S. B. Gelfand, the Editor for Transmission Systems of the IEEE Communications Society. Manuscript received April 15, 1999; revised January 30, 2001. This work was supported by the Independent Research Program of SSC San Diego, and by the ONR High-Data-Rate UHF Line-of-Sight Program.

M. Reuter, J. C. Allen, and R. C. North are with SSC San Diego, CA 92152 USA (e-mail: reuter@spawar.navy.mil).

J. R. Zeidler is with SSC San Diego, CA 92152 USA. He is also with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA.

Publisher Item Identifier S 0090-6778(01)10178-9.

Also, high-order modulation formats may be more susceptible than low-order formats. Therefore, a need exists to quantify DFE performance in the presence of decision error feedback and to mitigate the effects of error propagation.

Performance analysis of a DFE that includes decision error feedback is a difficult problem and has received considerable theoretical attention. Using an independence assumption on the noise and assuming no precursor ISI, Mosen used a Markov model of the error feedback process to derive probability of error for the DFE [1]. The mean burst error length also can be derived using this theory [8]. Although this approach is accurate, it becomes computationally burdensome because the number of Markov states exponentially increases as the number of feedback taps increases. Duttweiler, Mazo, and Messerschmitt simplified the analysis by aggregating the Markov states into a reduced-state model and then bounding the state transition probabilities [9]. Bounds on the probability of error [9] and the burst error length [10] are obtained. Tighter bounds are possible by refining the definition of the aggregate states [11], [12].

Various techniques for mitigating error propagation also have been proposed. Large taps in the feedback filter of an equalizer designed using the minimum mean-square-error (MSE) criterion cause significant self-generated ISI if decision errors are made. A potential solution is to jointly optimize the feedforward and feedback filters with an additional constraint on the norm of the feedback filter taps. This regularization approach was applied to the magnetic recording channel in [13], [14]. DFE structures also have been proposed recently that contain a soft decision device in the feedback path to compensate for unreliable decisions [15]–[18].

Combining these techniques, we present a minimum MSE DFE design that incorporates both a discrete soft decision device and a norm constraint on the feedback filter to mitigate error propagation. Reliable device inputs that are close with respect to the  $L^2$  norm to a component of the symbol constellation are fed back as hard decisions. Unreliable device inputs that are not close are fed back as intermediate decisions. Also, feedback gain is controlled by the norm constraint.

To capture the bursty nature of the DFE errors, we use both the probability of symbol error  $P_E$  and the mean burst error length  $\mu_{BL}$  as the primary DFE performance measures. We determine these quantities analytically using the fundamental approach of [1], but we derive an approximate Markov model with significantly fewer states by making an independence assumption involving the feedback error sequence. We demonstrate that design trade-offs between  $P_E$  and  $\mu_{BL}$  result in a classic “L-curve.” From this geometric interpretation, an optimum design becomes evident.

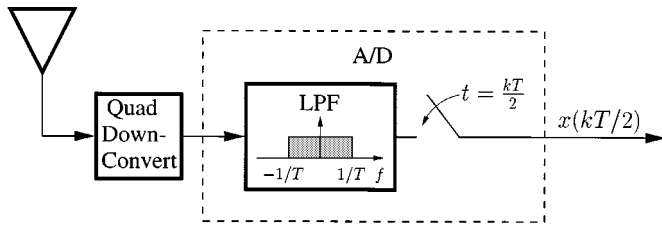


Fig. 1. Receiver front end.

## II. SYSTEM DEFINITIONS

We are concerned with coherent demodulation of data transmitted through a discrete multipath environment. The channel symbol modulation is  $M$ -ary phase shift keying ( $M$ -PSK). We operate on the complex baseband signal using the receiver front end shown in Fig. 1. The down-converter is locked to the carrier frequency  $f_c$  but not the phase. Therefore, any phase compensation is done by the equalizer [19]. The noise at the input to the ideal lowpass filter is modeled as a circularly symmetric complex white Gaussian process with two-sided power spectral density  $N_0$ .

The signal at the output of the ideal lowpass filter is modeled as

$$x(t) = \sum_{l=1}^L \beta_l(t) \sum_{m=-\infty}^{\infty} s(m) p_T(t - mT - \Delta_l - \tau) \times e^{-i2\pi f_c(\Delta_l + \tau)} + n(t) \quad (1)$$

where

- $L$  number of discrete multipath components;
- $\Delta_l$  relative path delays;
- $s(m)$  channel symbols;
- $T$  symbol duration.

The multiplicative fading terms  $\beta_l(t)$  incorporate all channel effects and are modeled as mutually independent random processes with zero mean and power  $\sigma_{\beta_l}^2$ . The corresponding path delays  $\Delta_l$  are deterministic. The symbol pulse  $p_T(t)$  is a square-root raised cosine waveform with unit energy. The variable  $\tau$  represents symbol timing error and is modeled as a continuous random variable uniformly distributed between  $-T/2$  and  $T/2$ . The  $M$ -PSK symbols  $s(m)$  are mutually independent random variables with unit power. The noise  $n(t)$  is a circularly symmetric Gaussian process with power  $\sigma_n^2 = 2N_0/T$ . Using these assumptions and by averaging over all random quantities, we define the received signal-to-noise ratio (SNR) to be [20]

$$\sigma_{\text{SNR}}^2 = \frac{1}{2N_0} \sum_{l=1}^L \sigma_{\beta_l}^2. \quad (2)$$

The fading on all paths is slow with respect to  $T$  so that we can assume a quasistatic channel with a given delay structure. All performance results are based upon the minimum MSE Wiener filter that is optimized with respect to a static realization of the channel defined by the set of parameters  $\{\beta_1, \beta_2, \dots, \beta_L, \tau\}$ . Also, the bandwidth of the lowpass filter in Fig. 1 accommodates sampling  $x(t)$  at the rate  $2/T$  to get the discrete samples  $x(kT/2)$ . This allows us to incorporate a fractionally spaced

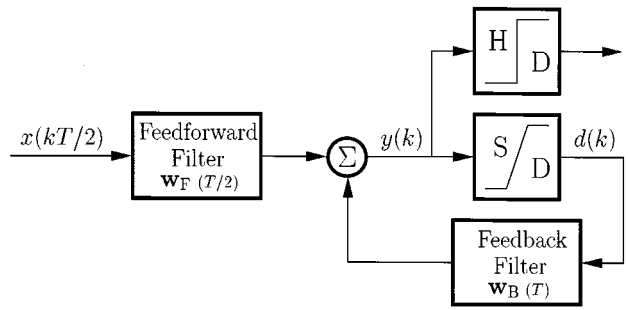


Fig. 2. DFE structure with soft decision feedback and hard decision output.

feedforward filter into the DFE design to provide robustness against timing error  $\tau$  as well as matched filtering to the transmitted pulse  $p_T(t)$  and the propagation channel [2].

## III. DFE DESIGN

A schematic of the DFE design is given in Fig. 2. The symmetric fractionally spaced feedforward filter  $\mathbf{w}_F$  has  $2N + 1$  taps with  $N$  taps leading and  $N$  taps lagging the cursor tap. The symbol-spaced feedback filter  $\mathbf{w}_B$  operating on the output of the soft decision device has  $N_B$  taps. We define the feedforward and feedback data vectors as

$$\begin{aligned} \mathbf{x}(k) &= [x(kT + NT/2) \quad x(kT + (N-1)T/2) \\ &\quad \dots \quad x(kT) \quad \dots \quad x(kT - NT/2)]^\dagger \\ \mathbf{d}(k) &= [d(k-1) \quad d(k-2) \quad \dots \quad d(k-N_B)]^\dagger \end{aligned} \quad (3)$$

where  $\dagger$  denotes the matrix transpose. Then the input to the hard decision device used to estimate the cursor symbol  $s(k)$  is written as

$$y(k) = \mathbf{w}_F^H \mathbf{x}(k) + \mathbf{w}_B^H \mathbf{d}(k) \quad (4)$$

where  $H$  denotes the Hermitian transpose.

There are many ways to define the soft decision device in Fig. 2. We propose the discrete implementation shown in Fig. 3 for 4-PSK. Decision regions for higher-order  $M$ -PSK scale accordingly. The shaded areas represent the intermediate decision regions and are defined by the angle  $\theta$ . The intermediate decisions are averages of the adjacent complex  $M$ -PSK symbols. The intent is to minimize the distance between the intermediate decision and the adjacent symbols. Applying this reasoning to 2-PSK modulation results in the erasure scheme presented in [16].

We derive the optimum constrained Wiener weights for the DFE by assuming error-free past decisions obtained from a training sequence so that  $d(k) = s(k)$ . The Wiener realization of the DFE is the solution of the regularization problem [21]

$$\begin{aligned} &\min \{E[\|s(k) - y(k)\|^2]\} \\ &\text{subject to } \|\mathbf{w}_B\|^2 = \gamma \|\mathbf{w}_{B_U}\|^2 \end{aligned} \quad (5)$$

where  $E[\cdot]$  is the expectation operator taken over the information and noise processes, and  $\|\mathbf{w}_{B_U}\|$  is the norm of the unconstrained feedback weights. The parameter  $\gamma$  controls the norm constraint where  $0 \leq \gamma \leq 1$ .

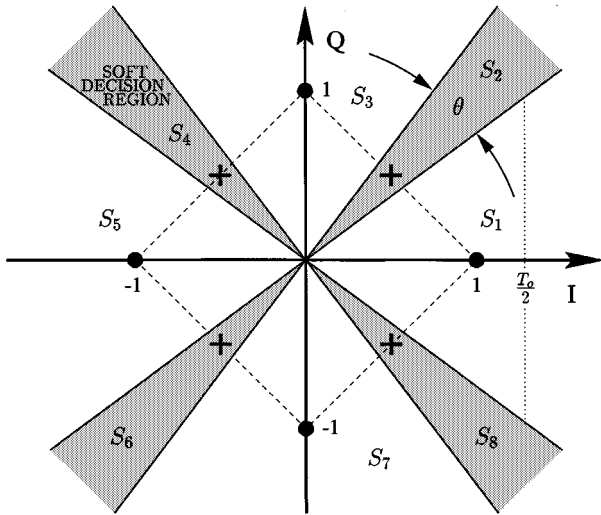


Fig. 3. Soft decision regions for 4-PSK with intermediate decisions in shaded regions.

Using the method of Lagrange multipliers, the optimum Wiener weights are readily shown to be the solution of the set of linear equations [2]

$$\begin{bmatrix} \mathbf{R} & \mathbf{Q}^H \\ \mathbf{Q} & (1 + \lambda)\mathbf{I}_{N_B} \end{bmatrix} \begin{bmatrix} \mathbf{w}_F \\ \mathbf{w}_B \end{bmatrix} = \begin{bmatrix} \mathbf{p} \\ \mathbf{0}_{N_B} \end{bmatrix} \quad (6)$$

where  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}(k)^H]$ ,  $\mathbf{Q} = E[\mathbf{d}(k)\mathbf{x}(k)^H]$ ,  $\mathbf{p} = E[s^*(k)\mathbf{x}(k)]$ , and  $*$  denotes complex conjugation.  $\mathbf{I}_{N_B}$  is the identity matrix of size  $N_B$ ,  $\mathbf{0}_{N_B}$  is the all-zero column vector of length  $N_B$ , and  $\lambda$  is the Lagrange multiplier  $\lambda \geq 0$ . Equation (6) shows that the norm constraint analytically introduces synthetic white noise with power  $\lambda$  into the feedback path.

The optimum feedforward and feedback filter weights are written explicitly as

$$\begin{aligned} \mathbf{w}_F &= (\mathbf{R} - \alpha\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{p} \\ \mathbf{w}_B &= -\alpha\mathbf{Q}\mathbf{w}_F \end{aligned} \quad (7)$$

where  $\alpha = 1/(1 + \lambda)$  and  $0 \leq \alpha \leq 1$ . If  $\alpha = 1$ , the optimum weights represent the standard unconstrained DFE. As  $\alpha \rightarrow 0$ , the feedback path diminishes and the feedforward filter  $\mathbf{w}_F$  approaches the optimum linear equalizer [13].

For convenience, we will define the norm constraint in terms of the parameter  $\gamma$  in (5). However, there is no closed-form relationship between  $\alpha$  and  $\gamma$  except at the extreme values where  $\gamma = 0$  and  $\gamma = 1$  correspond to  $\alpha = 0$  and  $\alpha = 1$ , respectively. Therefore, when invoking a constraint with arbitrary  $\gamma$ , we use (7) in a recursive search strategy to find the corresponding  $\alpha$ . Although we are not concerned with implementation issues, we note that if the DFE weights are adapted with the least mean square algorithm (LMS), the constrained equalizer can be implemented with only a minor increase in complexity using leaky LMS [13], [22, pp. 746–747].

#### IV. DFE PERFORMANCE ANALYSIS

We use the average probability of symbol error  $P_E$  and mean burst error length  $\mu_{BL}$  as the two primary measures of performance for the DFE. The approach we take to compute these two values is to separate the decision device input  $y(k)$  into components representing the self-generated ISI resulting from decision error feedback, the remaining residual ISI not spanned by the feedback filter  $\mathbf{w}_B$ , and noise. We further separate the error feedback sequence of length  $N_B$  into the actual transmitted sequence and a multiplicative process that represents potential  $M$ -PSK phase shift errors at the output of the soft decision device. We then make a simplifying assumption that this multiplicative error process, the transmitted symbols, and the noise process are statistically independent of each other. An approximate Markov model of the DFE follows in which each Markov state is defined by a permutation of the multiplicative error sequence. Using this independence assumption, we calculate components of the corresponding Markov probability transition matrix by averaging over the noise and all possible permutations of the transmitted  $M$ -PSK symbols contained within the sample  $y(k)$  using a computational technique presented in [23]. Then straightforward computations produce the approximate values  $P_E$  and  $\mu_{BL}$ .

The analysis begins by decomposing the feedforward data vector using (1) and (3) as

$$\mathbf{x}(k) = \sum_{l=1}^L \beta_l \bar{\mathbf{T}}_l \bar{\mathbf{s}}(k) + \mathbf{n}(k) \quad (8)$$

where  $\bar{\mathbf{s}}(k)$  is the infinite-length symbol vector

$$\bar{\mathbf{s}}(k) = [\cdots \quad s(k+1) \quad s(k) \quad s(k-1) \quad \cdots]^\dagger \quad (9)$$

and  $\mathbf{n}(k)$  is a  $(2N + 1) \times 1$  white complex Gaussian vector process with power  $\sigma_n^2$ . Each column of the  $(2N + 1) \times \infty$  channel matrix  $\bar{\mathbf{T}}_l$  is a replica of the fractionally sampled waveform  $p_T(t - \Delta_l - \tau) \exp(-i2\pi f_c(\Delta_l + \tau))$ , down-shifted by two samples with respect to the column on the left and truncated by the observation window of length  $2N + 1$ . By construction, the equalizer is coarse-synchronized to the path with  $\Delta_l = 0$  so that the center row of each  $\bar{\mathbf{T}}_l$  is given by

$$\begin{aligned} \bar{\mathbf{T}}_l(N + 1, :) \\ = [\cdots \quad p_T(T + \Delta_l + \tau) \quad p_T(\Delta_l + \tau) \\ p_T(-T + \Delta_l + \tau) \quad \cdots] e^{-i2\pi f_c(\Delta_l + \tau)}. \end{aligned} \quad (10)$$

Because the computational technique in [23] requires a finite number of ISI symbols, we truncate each  $\bar{\mathbf{T}}_l$  to get the  $(2N + 1) \times (2d + 1)$  matrix  $\mathbf{T}_l$  such that we maintain  $d$  columns on either side of the column containing the sample  $p_T(\Delta_l + \tau)$  corresponding to the cursor symbol  $s(k)$  in (10). Because of the fast roll-off of the square-root raised cosine pulse, the choice of the truncation parameter  $d$  is not critical, although it must be chosen with respect to the largest multipath delay such that  $d \gg \max\{|\Delta_l|\}/T$  and  $d > N_B$  ( $d = 30$  in all numerical results presented here).

Then by defining the composite channel matrix  $\mathbf{T}$  as

$$\mathbf{T} = \sum_{l=1}^L \beta_l \mathbf{T}_l \quad (11)$$

the feedforward data vector is well approximated by

$$\mathbf{x}(k) = \mathbf{T}\mathbf{s}(k) + \mathbf{n}(k), \quad (12)$$

where  $\mathbf{s}(k)$  is the truncated symbol vector given by

$$\mathbf{s}(k) = [s(k+d) \quad s(k+d-1) \quad \cdots \quad s(k) \quad \cdots \quad s(k-d)]^\dagger. \quad (13)$$

We next partition  $\mathbf{T}$  and  $\mathbf{s}(k)$  such that

$$\begin{aligned} \mathbf{T}\mathbf{s}(k) &= [\mathbf{E}_1 \quad \mathbf{f} \quad \mathbf{E}_2 \quad \mathbf{E}_3] \begin{bmatrix} \mathbf{s}_1(k) \\ s(k) \\ \mathbf{s}_2(k) \\ \mathbf{s}_3(k) \end{bmatrix} \\ &= \mathbf{E}_1 \mathbf{s}_1(k) + \mathbf{f} s(k) + \mathbf{E}_2 \mathbf{s}_2(k) + \mathbf{E}_3 \mathbf{s}_3(k) \end{aligned} \quad (14)$$

where

$$\begin{aligned} \mathbf{s}_1(k) &= [s(k+d) \quad s(k+d-1) \quad \cdots \quad s(k+1)]^\dagger, \\ \mathbf{s}_2(k) &= [s(k-1) \quad s(k-2) \quad \cdots \quad s(k-N_B)]^\dagger, \\ \mathbf{s}_3(k) &= [s(k-N_B-1) \quad s(k-N_B-2) \quad \cdots \quad s(k-d)]^\dagger. \end{aligned} \quad (15)$$

Then the cross-correlation components comprising the optimum Wiener weights in (7) are given by  $\mathbf{R} = \mathbf{T}\mathbf{T}^H + \sigma_n^2 \mathbf{I}_{2N+1}$ ,  $\mathbf{Q} = \mathbf{E}_2^H$ , and  $\mathbf{P} = \mathbf{f}$ . Also, using (4), (7), (12), and (14), we can write the decision device input as

$$\begin{aligned} y(k) &= \mathbf{w}_F^H (\mathbf{x}(k) - \alpha \mathbf{Q}^H \mathbf{d}(k)) \\ &= \mathbf{w}_F^H (\mathbf{T}\mathbf{s}(k) - \alpha \mathbf{E}_2 \mathbf{d}(k)) + \mathbf{w}_F^H \mathbf{n}(k) \\ &= \mathbf{w}_F^H (\mathbf{E}_1 \mathbf{s}_1(k) + \mathbf{f} s(k) + \mathbf{E}_3 \mathbf{s}_3(k)) \\ &\quad + \mathbf{w}_F^H \mathbf{E}_2 (\mathbf{s}_2(k) - \alpha \mathbf{d}(k)) + \mathbf{w}_F^H \mathbf{n}(k). \end{aligned} \quad (16)$$

The feedback soft decisions can be represented as

$$\mathbf{d}(k) = \mathbf{r}(k) \odot \mathbf{s}_2(k) \quad (17)$$

where  $\odot$  denotes element-wise product and  $\mathbf{r}(k)$  is a random  $N_B \times 1$  vector in which each term represents the rotation and scaling of the output of the soft decision device with respect to the transmitted  $M$ -PSK symbol (Fig. 3). If there are no decision errors,  $\mathbf{r}(k)$  is a vector of ones ( $\mathbf{r}(k) = \mathbf{1}_{N_B}$ ). From (16) we see that if the norm constraint is used ( $\alpha < 1$ ), and  $\mathbf{r}(k) = \mathbf{1}_{N_B}$ , the feedback filter does not cancel all of the postcursor symbols within its span.

The feedback error term in (16) then can be written as

$$\begin{aligned} \mathbf{s}_2(k) - \alpha \mathbf{d}(k) &= (\mathbf{1}_{N_B} - \alpha \mathbf{r}(k)) \odot \mathbf{s}_2(k) \\ &= \mathbf{e}(k) \odot \mathbf{s}_2(k) \end{aligned} \quad (18)$$

where  $\mathbf{e}(k)$  is the random multiplicative error vector. The possible values for each element of  $\mathbf{e}(k)$  are displayed in Fig. 4 for

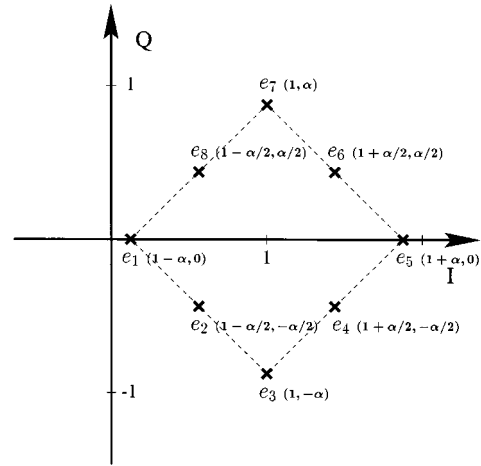


Fig. 4. Multiplicative feedback error values for 4-PSK with soft decisions ( $\theta > 0$ ) and a norm constraint ( $\alpha < 1$ ).

4-PSK using the soft decisions ( $\theta > 0$ ) given in Fig. 3. If  $\theta = 0$ , the intermediate (even) terms in Fig. 4 do not occur. The decision device input then can be decomposed in terms of the cursor symbol ISI and noise as

$$\begin{aligned} y(k) &= a_0 s(k) + \mathbf{a}_1 \mathbf{s}_1(k) + \mathbf{a}_2 (\mathbf{e}(k) \odot \mathbf{s}_2(k)) \\ &\quad + \mathbf{a}_3 \mathbf{s}_3(k) + \mathbf{w}_F^H \mathbf{n}(k) \end{aligned} \quad (19)$$

where  $a_0 = \mathbf{w}_F^H \mathbf{f}$ , and the remaining ISI coefficient row vectors are given by  $\mathbf{a}_j = \mathbf{w}_F^H \mathbf{E}_j$ .

The approach taken by Mosen [1] assumes the residual ISI involving the terms  $\mathbf{s}_1(k)$  and  $\mathbf{s}_3(k)$  in (19) is negligible. Then, making the reasonable approximation that the noise term is white Gaussian and therefore independent of the feedback error process, a Markov model of the DFE results in which each state is defined by a realization of the feedback error sequence  $\mathbf{e}(k) \odot \mathbf{s}_2(k)$ . By conditioning  $y(k)$  on the cursor symbol  $s(k)$  and each realization of  $\mathbf{e}(k) \odot \mathbf{s}_2(k)$ , a Gaussian variate results and is used to compute components of the transition probability matrix  $\mathbf{P}$ . Although we could take this approach by expanding the definition of each state to include realizations of the residual ISI, the number of states would be  $M^{2d-N_B} (2M^2)^{N_B}$  for  $M$ -PSK with soft decisions, and therefore computationally unwieldy.

The approach proposed in [24] accomplishes a significant reduction in states by assuming the residual ISI not spanned by the feedback filter is statistically independent of  $\mathbf{e}(k) \odot \mathbf{s}_2(k)$  and the noise. The states are defined as in [1] but the transition probabilities are computed by averaging over the residual ISI using an efficient numerical technique. Adapting this approximate method to  $M$ -PSK with soft decisions would result in  $(2M^2)^{N_B}$  states.

We achieve a further reduction in states (RIS) by making the following assumptions:

- RIS-1.** The multiplicative error vector  $\mathbf{e}(k)$  is statistically independent of the noise  $\mathbf{n}(k)$ .
- RIS-2.**  $\mathbf{e}(k)$  is independent of  $s(k)$  and the residual ISI represented by  $\mathbf{s}_1(k)$  and  $\mathbf{s}_3(k)$ .
- RIS-3.**  $\mathbf{e}(k)$  is independent of  $\mathbf{s}_2(k)$ , the symbol components spanned by the feedback filter.

Although not defined in terms of  $\mathbf{e}(k)$ , RIS-1 was used in [1] and RIS-2 was used in [24]. RIS-3 was used in [4] for 2-PSK, and amounts to lumping together in the same state all feedback error sequences  $\mathbf{e}(k) \odot \mathbf{s}_2(k)$  corresponding to the same  $\mathbf{e}(k)$  and assuming they have equal probability of occurring. In effect, because the distribution of the error sequences is unknown, we make the approximation that they are uniformly distributed. This is the key assumption in this paper and is a generalization of [4] to  $M$ -PSK. An approximate Markov model of the DFE follows from RIS-3 in which each of the  $(2M)^{N_B}$  states is defined by a realization of  $\mathbf{e}(k)$ .

To compute the state transition probabilities of the Markov matrix  $\mathbf{P}$ , we condition  $y(k)$  in (19) on  $\mathbf{e}(k)$  representing each state. The transition probabilities from the  $p$ th state are given by

$$\rho_p^q = \Pr(y(k) \in S_q | \mathbf{e}_p, s(k) = 1), \quad 1 \leq q \leq 2M \quad (20)$$

where  $\mathbf{e}_p$  is the multiplicative error sequence of the  $p$ th state, and  $S_q$  is the soft decision region in the complex plane corresponding to the transmitted symbol  $s(k) = 1$  and the multiplicative feedback error  $e_q$ . We construct the matrix such that the transition probabilities from the  $p$ th state reside in the  $p$ th row of  $\mathbf{P}$ . Also, the first row of  $\mathbf{P}$  corresponds to the error-free state in which each element of  $\mathbf{e}_1$  is  $e_1$  (Fig. 4).

Once we have constructed  $\mathbf{P}$ , we compute the steady-state probability vector  $\mathbf{v}$  as

$$\mathbf{P}^\dagger \mathbf{v} = \mathbf{v} \quad (21)$$

where the component  $v(p)$  represents the probability of being in the  $p$ th state [1]. Then the probability of  $M$ -PSK symbol error at the output of the hard decision device is [16]

$$P_E = \sum_{p=1}^{(2M)^{N_B}} v(p) \Pr(y(k) \notin H_1 | \mathbf{e}_p, s(k) = 1) \quad (22)$$

and the probability of bit error is

$$P_B = \frac{1}{\log_2 M} \sum_{p=1}^{(2M)^{N_B}} \sum_{m=2}^M \times v(p) W_m \Pr(y(k) \in H_m | \mathbf{e}_p, s(k) = 1) \quad (23)$$

where  $H_m$  is the hard decision region in the complex plane corresponding to the transmitted symbol  $s(k) = m$ , and  $W_m$  is the Hamming weight of the bits assigned to symbol  $m$  with respect to symbol  $m = 1$ . We calculate the probabilities in (20), (22), and (23) by averaging over the noise and all permutations of the symbol vectors  $\mathbf{s}_1(k)$ ,  $\mathbf{s}_2(k)$ , and  $\mathbf{s}_3(k)$  in (19) using a computational technique. Details are in the Appendix.

Next we partition the Markov matrix as

$$\mathbf{P} = \begin{bmatrix} \rho_1^1 & \mathbf{r}^\dagger \\ \mathbf{c} & \mathbf{G} \end{bmatrix}. \quad (24)$$

If we set  $\gamma = 1$  and  $\theta = 0$  ( $S_1 = H_1$ ), then the probability of symbol error assuming error-free feedback for the standard

TABLE I  
CHANNEL MODEL WITH PATH POWERS AND NORMALIZED DELAYS

Path	$\sigma_{\beta_l}^2$ (dB)	$\Delta_l/T$
1	-6.0	-0.8
2	0.0	0.0
3	-2.0	1.8
4	-4.0	2.7
5	-6.0	3.6

TABLE II  
STATIC CHANNEL REALIZATIONS

	Channel 1	Channel 2
$\tau/T$	0.3837	-0.4498
$\beta_1$	0.1916 + i0.3570	0.0346 + i0.3766
$\beta_2$	0.4836 + i0.2146	0.5033 - i0.2058
$\beta_3$	-0.3314 - i0.4589	-0.8192 + i0.4337
$\beta_4$	-0.1165 - i0.2191	0.4052 + i0.5288
$\beta_5$	0.5376 + i0.3074	-0.0024 + i0.1326

DFE is given by  $P_{NE} = 1 - \rho_1^1$ , and is the traditional optimum measure of DFE performance. By making the error-free state absorbing [8], we define the mean burst error length as [25, pp. 43–52]

$$\mu_{BL} = \boldsymbol{\pi}_0^\dagger \mathbf{N} \mathbf{1} - (N_B - 1) \quad (25)$$

where  $\mathbf{N} = (\mathbf{I} - \mathbf{G})^{-1}$ , and  $\boldsymbol{\pi}_0$  is the initial error state probability vector that we define as

$$\boldsymbol{\pi}_0 = \frac{\mathbf{r}}{1 - \rho_1^1}. \quad (26)$$

We include the term  $N_B - 1$  in the definition of  $\mu_{BL}$  so that the minimum mean burst error length is 1. Also, the variance of the burst error length is readily computed with this theory and is given by [8], [25, p. 52]

$$\sigma_{BL}^2 = \boldsymbol{\pi}_0^\dagger (2\mathbf{N} - \mathbf{I}) \mathbf{N} \mathbf{1} - (\boldsymbol{\pi}_0^\dagger \mathbf{N} \mathbf{1})^2. \quad (27)$$

Components of  $\boldsymbol{\pi}_0$  represent transition probabilities out of the error-free state weighted by the probability of an error. This definition of  $\boldsymbol{\pi}_0$  is convenient for high-order modulation and soft decision feedback because it takes into account that certain errors are much more likely than others, e.g.,  $e_2$  and  $e_8$  are more likely than  $e_4$  and  $e_6$  for 4-PSK in Fig. 4.

We also can further reduce the computational complexity of this approximate Markov DFE model by combining states containing unlikely errors using

**RIS-4.** All state probabilities  $\rho_p^q$  in (20) that transition to errors  $e_q$  with magnitude that exceed some threshold, are lumped into the worst-case error represented by  $e_{M+1}$ .

Consequently, Markov states defined by  $\mathbf{e}_p$  that differ only in elements corresponding to these unlikely errors are incorporated into an aggregate state with  $e_{M+1}$  in these error locations. For instance, for 4-PSK we might use only the errors  $\{e_1, e_2, e_3, e_5, e_7, e_8\}$  in Fig. 4 with  $e_4$  and  $e_6$  incorporated into  $e_5$ , resulting in  $6^{N_B}$  Markov states rather than  $8^{N_B}$ . We

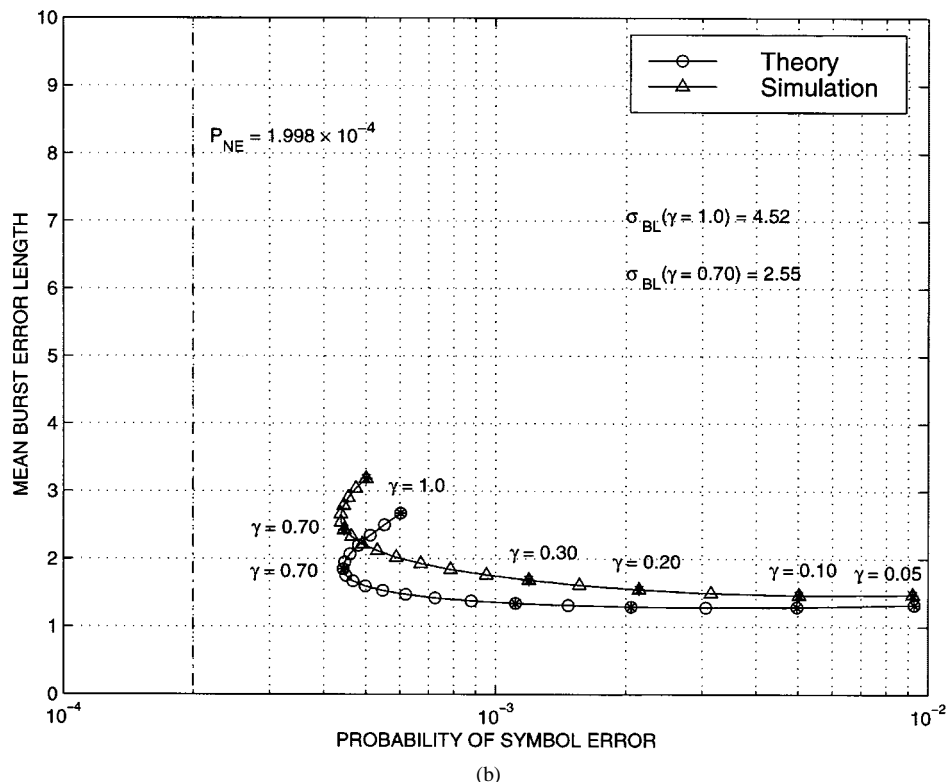
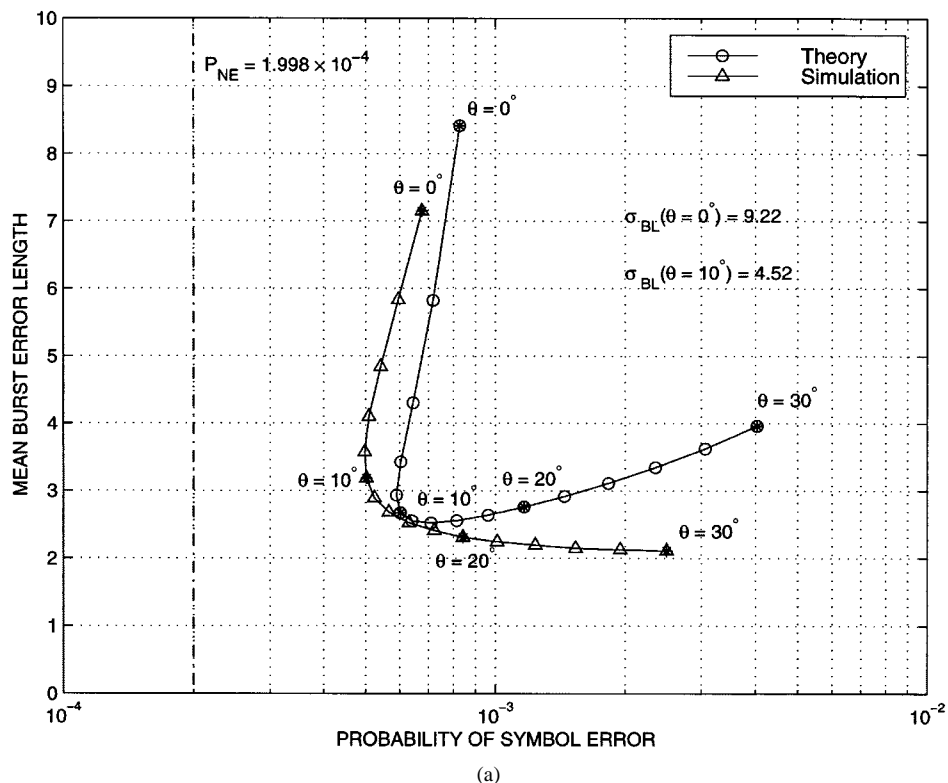


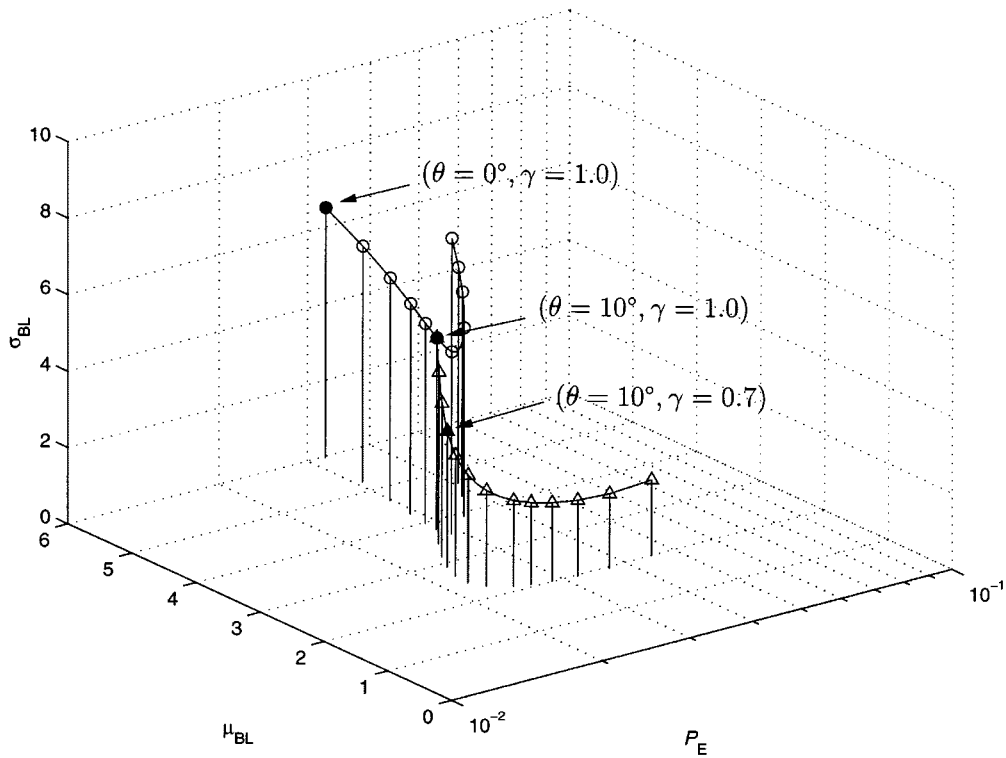
Fig. 5.  $\mu_{BL}$  versus  $P_E$  for Channel 1 with 4-PSK, and SNR = 15 dB. (a) Over soft decision angle  $\theta$  with  $\gamma = 1.0$ . (b) Over norm constraint  $\gamma$  with  $\theta = 10^\circ$ .

set the threshold such that the decrease in computational complexity is worth the potential loss in accuracy.

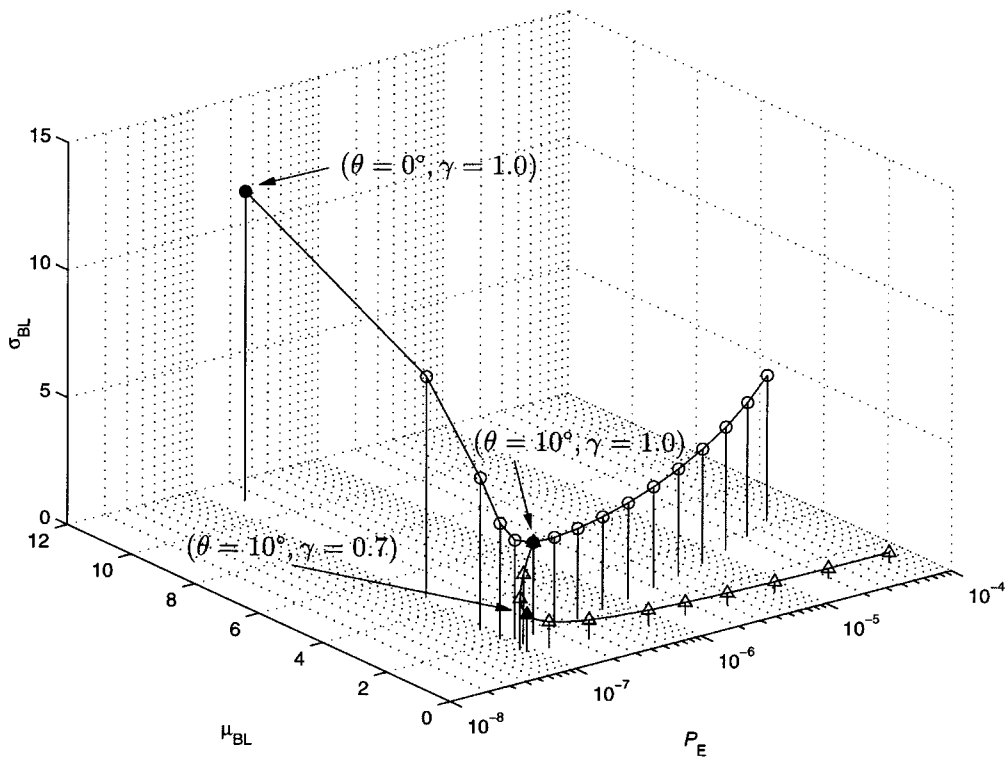
### V. NUMERICAL RESULTS

To demonstrate the utility of the DFE design and the performance analysis approach, we present results for two static

channel realizations derived from the channel model given in Table I. The fading parameters are Rayleigh distributed with given powers. The corresponding normalized path delays are also shown. The equalizer is coarse-synchronized to the second path. The two static channels extracted from this model are given in Table II, and present different degrees of challenge to



(a)



(b)

Fig. 6. Theoretical performance curves over soft decision angle ( $\theta$ ) and norm constraint ( $\gamma$ ) for Channel 1 with 4-PSK. (a) SNR = 11 dB. (b) SNR = 19 dB.

the DFE. The normalized timing error is also given. The DFE design consists of a fractionally spaced feedforward filter  $\mathbf{w}_F$  with  $N = 10$  pre- and postcursor taps, and a feedback filter  $\mathbf{w}_B$  with  $N_B = 4$  taps. The transmitted square-root raised cosine pulse has 30% excess bandwidth.

#### A. Channel 1: 4-PSK

We begin by equalizing Channel 1 with 4-PSK modulation and  $\sigma_{\text{SNR}}^2 = 15$  dB. Fig. 5(a) displays mean burst error length  $\mu_{\text{BL}}$  versus probability of symbol error  $P_E$  as the soft decision angle  $\theta$  varies from  $\theta = 0^\circ$  to  $\theta = 30^\circ$  in two-degree incre-

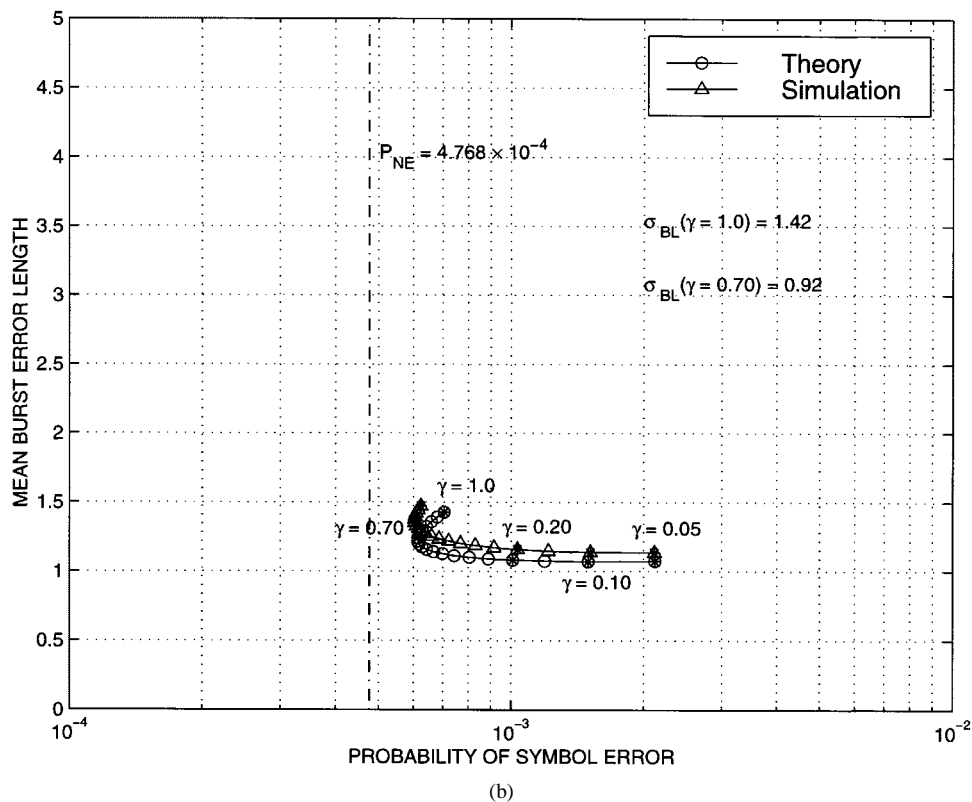
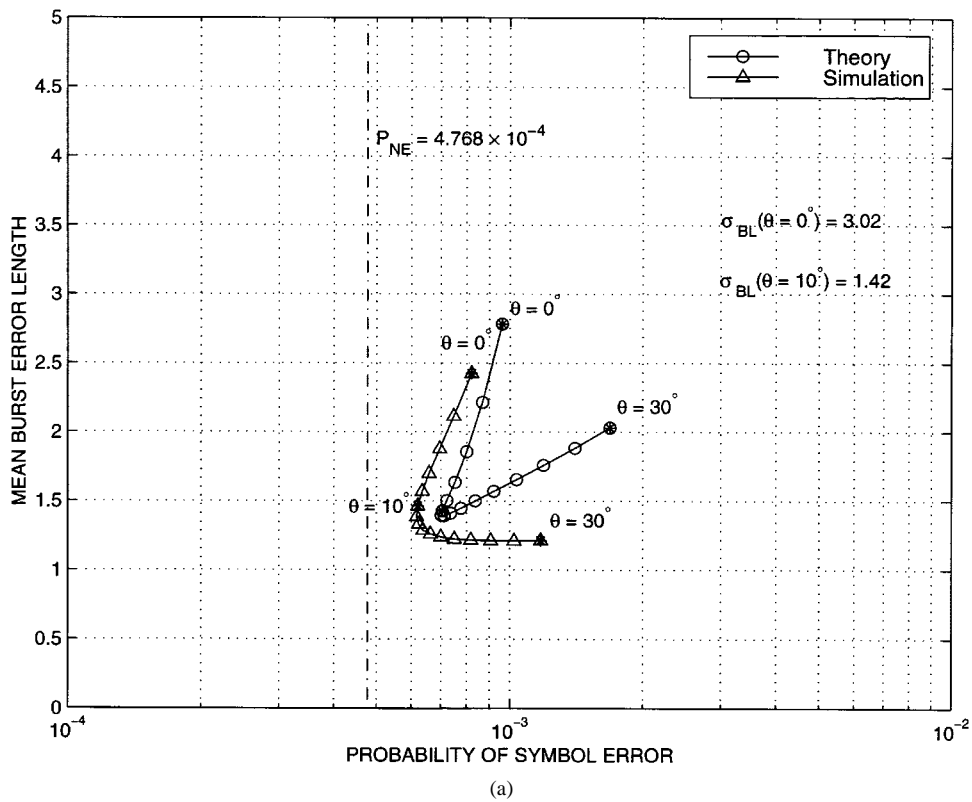


Fig. 7.  $\mu_{BL}$  versus  $P_E$  for Channel 2 with 4-PSK, and SNR = 12 dB. (a) Over soft decision angle  $\theta$  with  $\gamma = 1.0$ . (b) Over norm constraint  $\gamma$  with  $\theta = 10^\circ$ .

ments. No norm constraint is used ( $\gamma = 1$ ). Also shown is the probability of error of the standard unconstrained DFE given no feedback error  $P_{NE}$ . Theoretical and simulation results taken over  $2 \times 10^8$  transmitted symbols are displayed. In the simulations, error bursts are determined at the output of the hard de-

cision device in Fig. 2 and are defined to contain no more than  $N_B - 1$  consecutive correct decisions. However, the theoretical definition of  $\mu_{BL}$  in (25) is defined with respect to the soft decision device. And because an adjacent soft decision error ( $e_2$  and  $e_8$  in Fig. 4) may not necessarily trigger a hard decision error,

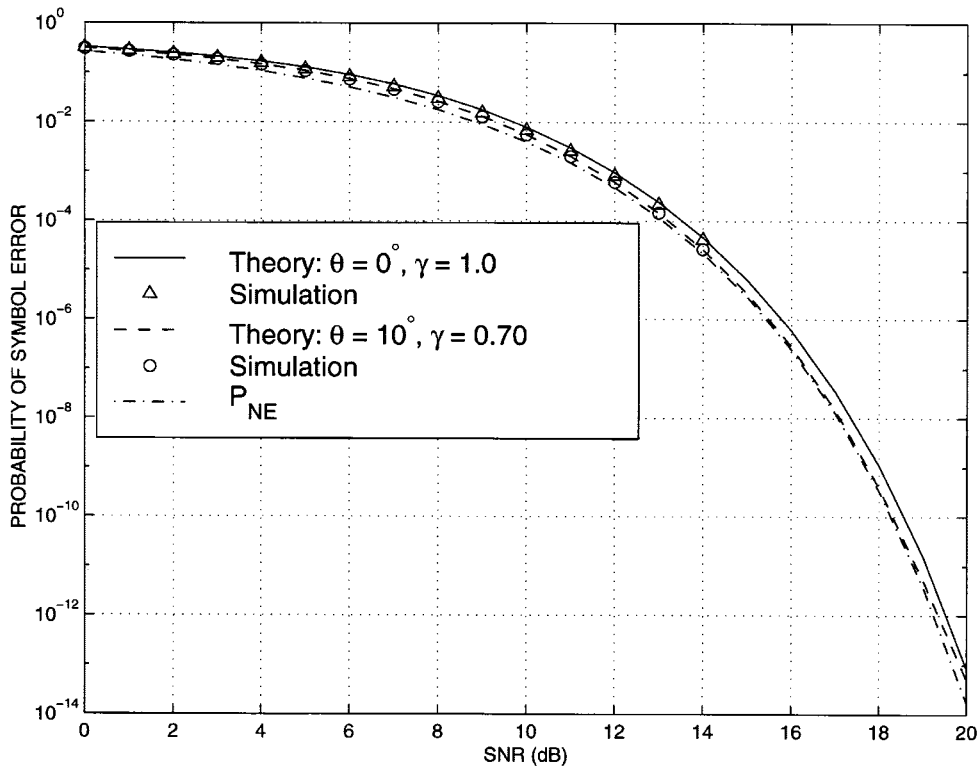


Fig. 8.  $P_E$  versus SNR for Channel 2 with 4-PSK.

theoretical and measured  $\mu_{BL}$  tend to deviate for large values of  $\theta$ .

Fig. 5(a) reveals that the traditional DFE using no soft decisions ( $\theta = 0^\circ$ ), suffers degradation in  $P_E$  with respect to the error-free feedback case  $P_{NE}$ . However, as we increase  $\theta$ , an “L-curve” [21] develops in which there is a decrease in  $P_E$  that is accompanied by a significant decrease in  $\mu_{BL}$  until  $\theta \approx 10^\circ$  when  $P_E$  begins to increase. Also shown is the considerable improvement in standard deviation of the burst error length  $\sigma_{BL}$  determined theoretically from (27) for  $\theta = 0^\circ$  and  $\theta = 10^\circ$ .

Next, by setting  $\theta = 10^\circ$ , we begin at the knee of the L-curve in Fig. 5(a) and plot  $\mu_{BL}$  versus  $P_E$  over the norm constraint  $\gamma$  in Fig. 5(b). We see that further improvement in  $P_E$  and  $\mu_{BL}$  is obtained by incorporating the norm constraint with the soft decision device. Again an L-curve develops in which both  $P_E$  and  $\mu_{BL}$  decrease until  $\gamma \approx 0.70$ . After this point, any decrease in  $\mu_{BL}$  comes at the expense of increasing  $P_E$ . At the operating point  $\gamma = 0.70$ , we have decreased  $P_E$  approximately by half with respect to the standard DFE, and have come closer to the optimum  $P_{NE}$ . We have also dramatically decreased  $\mu_{BL}$  and  $\sigma_{BL}$ . For the purpose of comparison, estimated  $\mu_{BL}$  and  $\sigma_{BL}$  derived from the simulations are  $\hat{\mu}_{BL} = 7.1$  and  $\hat{\sigma}_{BL} = 6.9$  for the parameter set ( $\theta = 0^\circ, \gamma = 1.0$ ), and  $\hat{\mu}_{BL} = 2.4$  and  $\hat{\sigma}_{BL} = 3.2$  for ( $\theta = 10^\circ, \gamma = 0.70$ ). We refer to the knee of this L-curve as the “optimum” operating point. Also, we achieve the same result if we generate the L-curves first over  $\gamma$  and then over  $\theta$ .

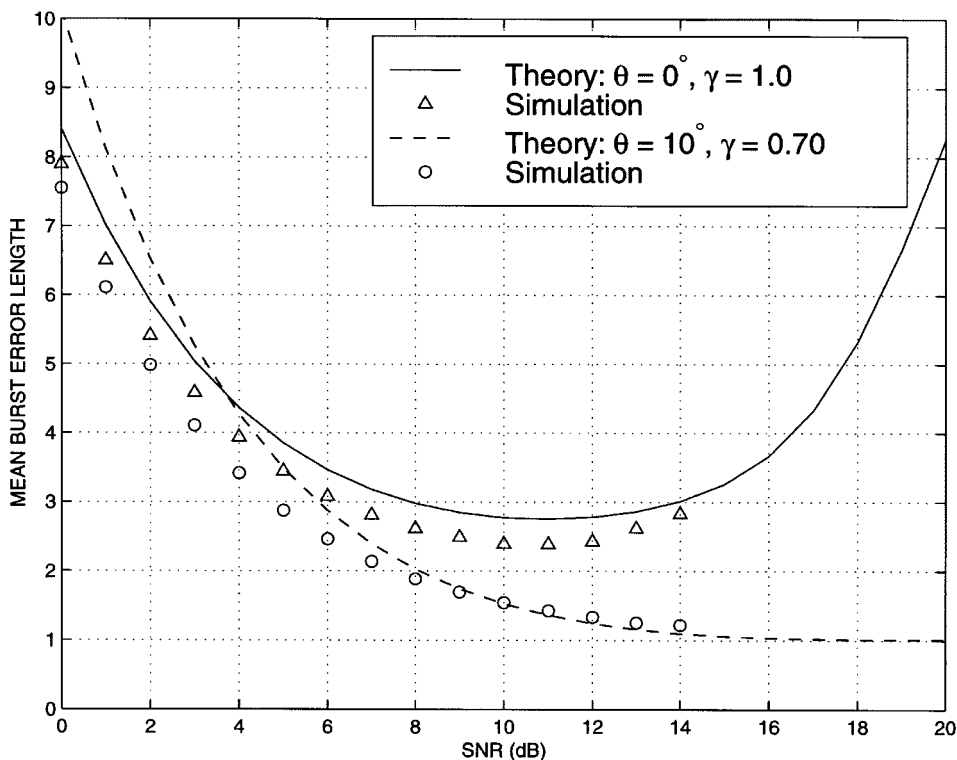
The “optimum” set of parameters  $(\theta, \gamma)$  generally is dependent on SNR. To demonstrate, Fig. 6 contains three dimensional plots of theoretical DFE performance for Channel 1 with  $\sigma_{SNR}^2 = 11$  dB and  $\sigma_{SNR}^2 = 19$  dB. The standard deviation  $\sigma_{BL}$

is plotted on the  $z$ -axis. In each plot, the performance curve over  $\theta$  with  $\gamma = 1$  is combined with the curve over  $\gamma$  with  $\theta = 10^\circ$ . These plots reveal that although we can achieve better performance at the low SNR using a different parameter set, the exact choice of  $\theta$  and  $\gamma$  is not critical because a range of parameters produces approximately the same performance. However, the set ( $\theta = 10^\circ, \gamma = 0.70$ ) represents a good compromise between the low and high SNRs. These plots also demonstrate the substantial improvement in  $\sigma_{BL}$  that can be obtained using the constrained DFE with soft decision feedback, particularly at the high SNRs.

### B. Channel 2: 4-PSK

Fig. 7 contains results for the more benign Channel 2 with  $\sigma_{SNR}^2 = 12$  dB. Although the performance gains are more modest, we still see the characteristic L-curve in both plots. Also, it is interesting that ( $\theta = 10^\circ, \gamma = 0.70$ ) is close to the optimum operating region of the DFE for both channels as indicated by Figs. 5(b) and 7(b). This was observed over a variety of channel realizations for 4-PSK.

DFE performance is given as a function of SNR for the two operating regions ( $\theta = 0^\circ, \gamma = 1.0$ ) and ( $\theta = 10^\circ, \gamma = 0.70$ ) in Figs. 8 and 9 for Channel 2 (similar effects are seen for Channel 1). Increasing SNR from 0 dB has the effect of improving  $P_E$  and  $\mu_{BL}$  for both DFE designs because in this region noise is the major cause of error bursts rather than decision error feedback. However, after approximately 11 dB, the mean burst error length of the standard DFE in Fig. 9 begins to increase even as  $P_E$  in Fig. 8 decreases. At this point the norm of the feedback filter increases to the extent that decision errors

Fig. 9.  $\mu_{BL}$  versus SNR for Channel 2 with 4-PSK.TABLE III  
BURST LENGTH (BL) SIMULATION STATISTICS, CHANNEL 2, 4-PSK, SNR = 15 dB

DFE Type	Number of Bursts					Maximum BL
	BL = 1	BL = 2	BL > 5	BL > 10	BL > 15	
$\gamma = 1, \theta = 0$	1356	413	517	183	61	36
$\gamma = .7, \theta = 10$	2945	53	29	3	0	12

begin to dominate. As SNR increases beyond 11 dB, initial decision errors tend to cause longer and longer strings of errors. However, bursts occur less frequently as the SNR increases to the extent that the longer bursts do not severely degrade the average probability of error. Conversely,  $\mu_{BL}$  of the constrained DFE with the soft decision device continues to decrease as SNR increases. To further illustrate, Table III contains burst length simulation statistics for SNR = 15 dB, measured over  $10^9$  transmitted channel symbols. Although the constrained DFE exhibits about twice as many single errors as the standard equalizer, there is a significant decrease in the number of bursts of length 2 and greater. Figs. 8 and 9 demonstrate the importance of including burst error statistics in the analysis of a DFE.

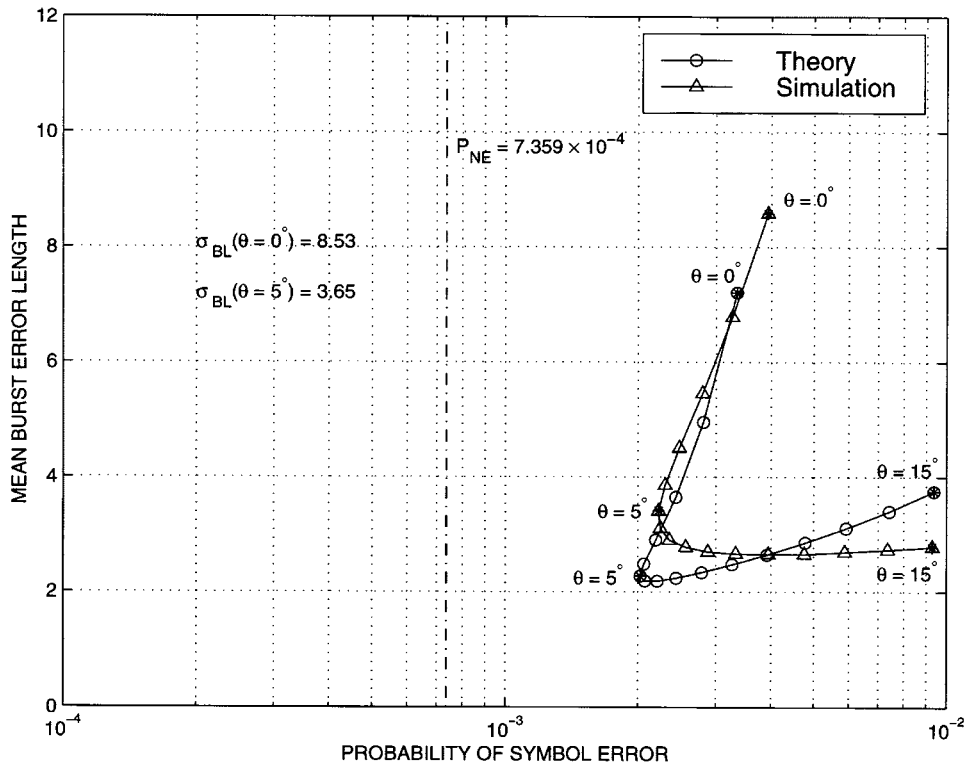
### C. Channel 2: 8-PSK

This example demonstrates DFE performance with 8-PSK modulation for Channel 2. Fig. 10 displays  $\mu_{BL}$  versus  $P_E$  in which we vary  $\theta$  and  $\gamma$  with  $\sigma_{SNR}^2 = 18$  dB. To minimize computations, we use the aggregate-state approach described by RIS-4 at the end of Section IV. Adapting the sequencing from 4-PSK in Fig. 4 to 8-PSK, we use the errors  $\{e_1, e_2, e_3, e_4, e_9, e_{14}, e_{15}, e_{16}\}$  to define the Markov states, with the omitted errors lumped into  $e_9$ . We then have  $8^{N_B}$

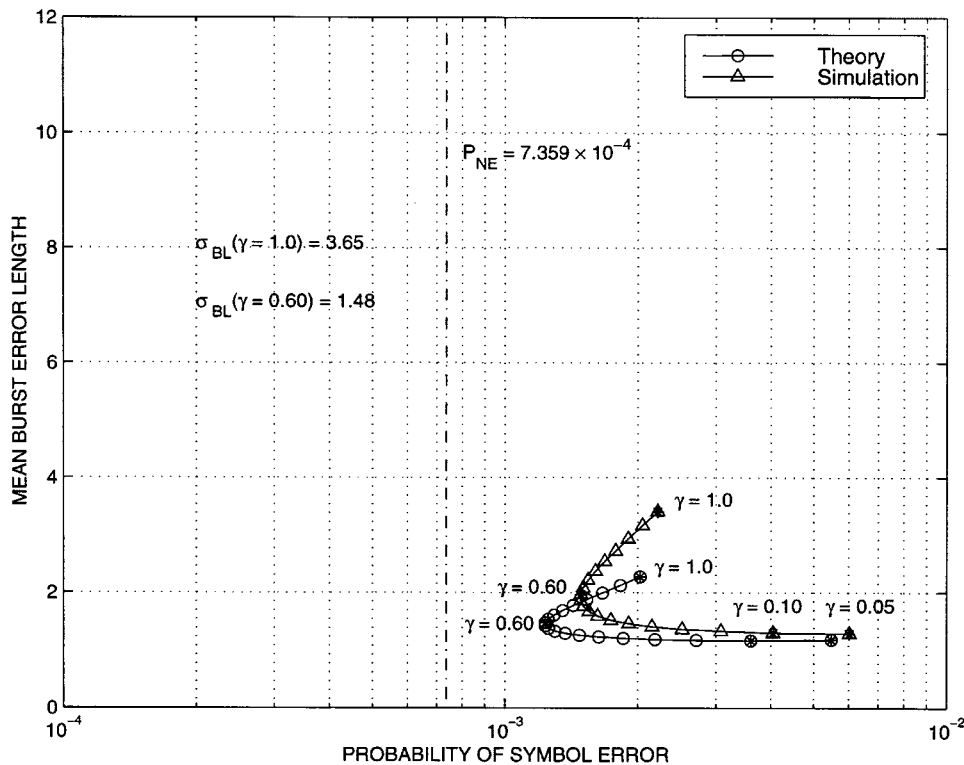
states, which is the same number used to generate the performance results for 4-PSK. Comparing Figs. 7 and 10 reveals that we need to increase the SNR for 8-PSK by more than 6 dB to achieve approximately the  $P_E$  performance of 4-PSK. Also, the burst error statistics  $\mu_{BL}$  and  $\sigma_{BL}$  have degraded significantly for the standard equalizer with ( $\theta = 0^\circ, \gamma = 1$ ) in Fig. 10(a). And yet, by using soft decisions and invoking the norm constraint, we see from Fig. 10(b) that we can get  $\mu_{BL}$  and  $\sigma_{BL}$  values comparable to those given in Fig. 7(b) for 4-PSK. Also, given the results for 4-PSK, it is intuitively satisfying that the optimum 8-PSK decision angle is  $\theta = 5^\circ$ .

### D. Channel 2: 4-PSK With Convolutional Coding

Although the increase in  $P_E$  of the standard DFE with respect to  $P_{NE}$  may appear minor in Fig. 8, the associated burst errors have significant implications for error correction coding and interleaving [6]. This is particularly true when operating at high SNR. To demonstrate, we incorporate into the design a simple rate 1/2 convolutional code with constraint length 3 and free distance 5 [26, p. 466] with hard-decision decoding. The channel symbols are 4-PSK with Gray bit mapping. An  $(N_i, N_i)$  convolutional interleaver is included between the output of the encoder and the 4-PSK channel symbol mapper, incurring an end-to-end delay of  $N_i(N_i - 1)$  [27, pp. 347–349].



(a)



(b)

Fig. 10.  $\mu_{BL}$  versus  $P_E$  for Channel 2 with 8-PSK, and SNR = 18 dB. (a) Over soft decision angle  $\theta$  with  $\gamma = 1.0$ . (b) Over norm constraint  $\gamma$  with  $\theta = 5^\circ$ .

Fig. 11 displays the probability of information bit error  $P_b$  at the output of the convolutional decoder as a function of SNR for Channel 2. The theoretical upper bound on the error of the constrained DFE ( $\theta = 10^\circ$ ,  $\gamma = 0.70$ ) assuming ideal bit interleaving is plotted. This bound is given by  $P_b < \sum_{d=5}^{\infty} \beta_d P_2(d)$ ,

where  $\beta_d = (d - 4)2^{d-5}$ , and the path error probability  $P_2(d)$  is determined using the channel bit error probability of (23) [26, pp. 463–466]. Also included are simulation results for the standard DFE and the constrained DFE with convolutional interleaving and  $N_i = 11$ . Although the constrained DFE mitigates

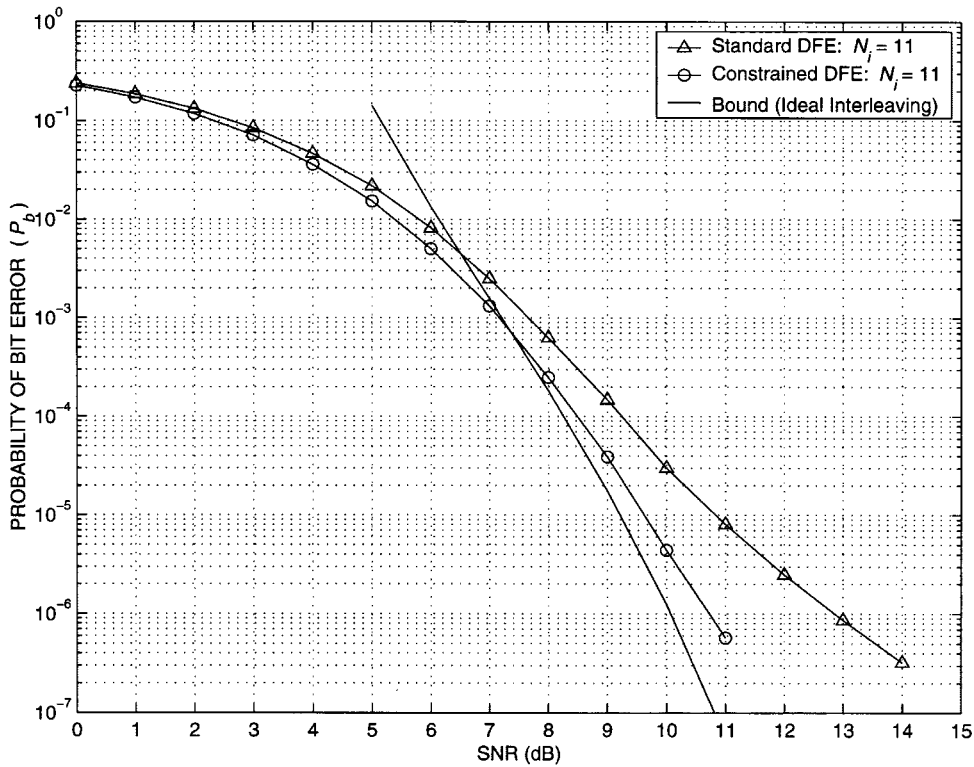


Fig. 11.  $P_b$  versus SNR for Channel 2 with rate 1/2 convolutional code, convolutional interleaving, and 4-PSK channel symbols.

burst errors, it does not eliminate them. And because convolutional codes are sensitive to burst errors, an interleaver is still required to achieve the modest coding gain of this code. However, for  $P_b = 10^{-6}$ , approximately 2-dB improvement in SNR is obtained with the constrained DFE versus the standard DFE. To achieve this same performance,  $N_i$  must double for the standard DFE, resulting approximately in a four-fold increase in end-to-end delay.

## VI. CONCLUSION

We have demonstrated that a DFE design that combines discrete soft decision feedback with a norm constraint on the feedback filter is an effective method of controlling error propagation with minor computational cost. We analyzed this structure using an approximate Markov model that predicts DFE performance in terms of probability of symbol error as well as burst decision error statistics and applied it to  $M$ -PSK modulation.

Other soft decision implementations than the simple one proposed here for  $M$ -PSK may offer further improvements in performance. For instance, we can quantize the decision space more finely or include an erasure region as in [16]. The applicability of the analysis approach to quadrature amplitude modulation (QAM) can be explored. Also, we may achieve better performance by using more complicated DFE designs that incorporate interleaving and error correction coding into the feedback path [28], [29]. Finally, we have considered only individual static channel realizations drawn from a fading propagation environment. Implementation issues related to an adaptive DFE operating in a time-varying environment and the effect of feedback errors on tracking performance warrant further analysis.

## APPENDIX PROBABILITY COMPUTATIONS

We calculate the probabilities in (20), (22), and (23) using a computationally efficient technique. It can be shown that [23]

$$\begin{aligned} & \Pr(y(k) \in D | \mathbf{e}_p, s(k) = 1) \\ &= \frac{1}{T_o^2} \sum_{l_Q = -P}^P \sum_{l_I = -P}^P \hat{K}_D(\omega l_I, \omega l_Q) h_{\text{ISI}}(\omega l_I, \omega l_Q) \\ & \quad \times h_N(\omega l_I, \omega l_Q) e^{i(\omega l_I \text{Re}\{a_0\} + \omega l_Q \text{Im}\{a_0\})} + \epsilon \end{aligned} \quad (28)$$

where

- $D$  region in the complex plane corresponding to either hard or soft  $M$ -PSK decisions;
- $\hat{K}_D(\omega l_I, \omega l_Q)$  Fourier transform of the wedge-shaped region  $D$  truncated at  $T_o/2$  (see Fig. 3 for  $D = S_1$ );
- $h_{\text{ISI}}(\omega l_I, \omega l_Q)$  characteristic function of the  $M$ -PSK ISI;
- $h_N(\omega l_I, \omega l_Q)$  characteristic function of the noise process  $\mathbf{w}_F^H \mathbf{n}(k)$ , and  $\omega = 2\pi/T_o$ .

The term  $\epsilon$  is the symbol error associated with this method and can be chosen arbitrarily small with the proper choice of the parameters  $T_o$  and  $P$  [23].

The characteristic function of the ISI can be decomposed according to the symbol vectors  $\mathbf{s}_1(k)$ ,  $\mathbf{s}_2(k)$ ,  $\mathbf{s}_3(k)$ , and the associated coefficients from (19) as

$$\begin{aligned} h_{\text{ISI}}(\omega l_I, \omega l_Q) &= \prod_{j=1}^d h_{1,j}(\omega l_I, \omega l_Q) \prod_{j=1}^{N_B} h_{2,j}(\omega l_I, \omega l_Q) \\ & \quad \times \prod_{j=1}^{d-N_B} h_{3,j}(\omega l_I, \omega l_Q). \end{aligned} \quad (29)$$

Each component of the ISI corresponding to  $\mathbf{s}_1(k)$  and  $\mathbf{s}_3(k)$  is given by

$$\begin{aligned} & h_{n,j}(\omega l_I, \omega l_Q) \\ &= \frac{2}{M} \sum_{m=1}^{M/2} \cos\left(\omega(\operatorname{Re}\{a_{n,j}\}l_I + \operatorname{Im}\{a_{n,j}\}l_Q)\cos\left(\frac{2\pi}{M}(m-1)\right)\right. \\ & \quad \left. + \omega(\operatorname{Re}\{a_{n,j}\}l_Q - \operatorname{Im}\{a_{n,j}\}l_I)\sin\left(\frac{2\pi}{M}(m-1)\right)\right) \end{aligned} \quad (30)$$

where  $\operatorname{Re}\{a_{n,j}\}$  and  $\operatorname{Im}\{a_{n,j}\}$  are the real and imaginary parts of the  $j$ th element of  $\mathbf{a}_n$  for  $n = 1, 3$ . The remaining error feedback terms are given by

$$\begin{aligned} & h_{2,j}(\omega l_I, \omega l_Q) \\ &= \frac{2}{M} \sum_{m=1}^{M/2} \cos\left(\omega(\operatorname{Re}\{a_{2,j}e_{p,j}\}l_I + \operatorname{Im}\{a_{2,j}e_{p,j}\}l_Q)\right. \\ & \quad \times \cos\left(\frac{2\pi}{M}(m-1)\right) \\ & \quad \left. + \omega(\operatorname{Re}\{a_{2,j}e_{p,j}\}l_Q - \operatorname{Im}\{a_{2,j}e_{p,j}\}l_I)\right. \\ & \quad \left. \times \sin\left(\frac{2\pi}{M}(m-1)\right)\right) \end{aligned} \quad (31)$$

where  $e_{p,j}$  is the  $j$ th element of the multiplicative error vector  $\mathbf{e}_p$  corresponding to the  $p$ th state. The characteristic function of the circular Gaussian noise is given by

$$h_N(\omega l_I, \omega l_Q) = e^{-\omega^2(l_I^2 + l_Q^2)\sigma^2/2} \quad (32)$$

where  $2\sigma^2$  is the variance of the complex process  $\mathbf{w}_F^H \mathbf{n}(k)$ .

The Fourier transform of a truncated wedge that is symmetric about the real axis is given by

$$\begin{aligned} \hat{K}_\phi(\omega l_I, \omega l_Q) &= \frac{i\pi}{\omega^2 l_Q} (-i)^{l_I} \left( \operatorname{sinc}\left(\frac{1}{2}(l_I + l_Q \tan(\phi))\right) \right. \\ & \quad \times e^{-i\pi l_Q \tan(\phi)/2} \\ & \quad \left. - \operatorname{sinc}\left(\frac{1}{2}(l_I - l_Q \tan(\phi))\right) e^{i\pi l_Q \tan(\phi)/2} \right), \\ & \quad l_Q \neq 0 \\ \hat{K}_\phi(0, 0) &= \tan(\phi) \frac{\pi^2}{\omega^2} \\ \hat{K}_\phi(\omega l_I, 0) &= \frac{i4 \tan(\phi)(-i)^{l_I}}{\omega^2 l_I^2} \left( \frac{\pi l_I}{2} (-i)^{l_I} - \sin\left(\frac{\pi}{2} l_I\right) \right) \\ & \quad l_I \neq 0 \end{aligned} \quad (33)$$

where  $\phi$  is the angle of the side of the wedge with respect to the real axis and  $\operatorname{sinc}(x) = \sin(\pi x)/\pi x$ .

For example, to determine the components of (22) for 4-PSK, we calculate

$$\begin{aligned} \Pr(y(k) \notin H_1 | \mathbf{e}_p, s(k) = 1) \\ = 1 - \Pr(y(k) \in H_1 | \mathbf{e}_p, s(k) = 1) \end{aligned} \quad (34)$$

using (28) and (33) with  $\phi = \pi/4$ . The Markov transition probabilities of (20) are determined similarly using (33) and the relation

$$\begin{aligned} \hat{K}_{S_q}(\omega l_I, \omega l_Q) &= \hat{K}_\phi(\omega(l_I \cos(\phi_q) + l_Q \sin(\phi_q)), \\ & \quad \omega(l_Q \cos(\phi_q) - l_I \sin(\phi_q))) \end{aligned} \quad (35)$$

where  $\phi_q$  is the angle of rotation of the soft decision region  $S_q$  with respect to the real axis. For the 4-PSK soft decisions defined in Fig. 3, we use the wedge angles  $\phi \in \{\theta/2, \pi/4 - \theta/2\}$  in (33).

## REFERENCES

- [1] P. Mosen, "Adaptive equalization of the slow fading channel," *IEEE Trans. Commun.*, vol. COM-22, pp. 1064–1075, Aug. 1974.
- [2] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.
- [3] P. Mosen, "MMSE equalization of interference on fading diversity channels," *IEEE Trans. Commun.*, vol. COM-32, pp. 5–12, Jan. 1984.
- [4] L. Li and L. B. Milstein, "Rejection of CW interference in QPSK systems using decision-feedback filters," *IEEE Trans. Commun.*, vol. Com-31, pp. 473–483, Apr. 1983.
- [5] K. A. Hamied and G. L. Stüber, "Performance of trellis-coded modulation for equalized multipath fading ISI channels," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 50–58, Feb. 1995.
- [6] D. J. Costello, P. R. Chevillat, and L. C. Perez, "Channel coding considerations for wireless LAN's," in *Proc. IEEE Int. Symp. Inform. Theory*, 1997, p. 462.
- [7] P. Mosen, "Theoretical and measured performance of a DFE modem on a fading multipath channel," *IEEE Trans. Commun.*, vol. COM-25, pp. 1144–1153, Oct. 1977.
- [8] R. A. Kennedy and B. D. O. Anderson, "Recovery times of decision feedback equalizers on noiseless channels," *IEEE Trans. Commun.*, vol. COM-35, pp. 1012–1021, Oct. 1987.
- [9] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 490–497, July 1974.
- [10] N. C. Beaulieu, "Bounds on recovery times of decision feedback equalizers," *IEEE Trans. Commun.*, vol. 42, pp. 2786–2794, Oct. 1994.
- [11] S. A. Altekar and N. C. Beaulieu, "Upper bounds to the error probability of decision feedback equalization," *IEEE Trans. Inform. Theory*, vol. 39, pp. 145–156, Jan. 1993.
- [12] W. W. Choy and N. C. Beaulieu, "Improved bounds for error recovery times of decision feedback equalization," *IEEE Trans. Inform. Theory*, vol. 43, pp. 890–902, May 1997.
- [13] P. S. Bednarz and J. M. Cioffi, "Decision feedback equalization for channels with error correcting capabilities," in *Proc. IEEE Int. Conf. Commun. (ICC '97)*, 1997, pp. 1607–1612.
- [14] G. Mathew, B. Farhang-Boroujeny, R. W. Wood, and L. Bin, "Constrained equalizer design for MDFE detection on the magnetic recording channel," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM 97)*, 1997, pp. 1258–1262.
- [15] R. A. Kennedy and Z. Ding, "Quantizer design and optimization in decision feedback equalization," in *Int. Symp. Signal Processing and its Applications (ISSPA-96)*, 1996, pp. 614–617.
- [16] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 757–760, July 1997.
- [17] Y. H. Kim and S. Shamsunder, "Adaptive algorithms for channel equalization with soft decision feedback," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1660–1669, Dec. 1998.
- [18] S. L. Ariyavisitakul and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1670–1678, Dec. 1998.
- [19] J. R. Treichler, I. Fijalkow, and C. R. Johnson, "Fractionally spaced equalizers," *IEEE Signal Processing Mag.*, vol. 13, pp. 65–81, May 1996.
- [20] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio—Part I: Theoretical considerations," *IEEE Trans. Commun.*, vol. 40, pp. 885–894, May 1992.
- [21] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Rev.*, vol. 34, no. 4, pp. 561–580, Dec. 1992.
- [22] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [23] M. Reuter, "Numerically efficient Fourier-based technique for calculating error probabilities with intersymbol interference," *IEEE Trans. Commun.*, vol. 45, pp. 629–632, June 1997.
- [24] J. E. Smee and N. C. Beaulieu, "Error-rate evaluation of linear equalization and decision feedback equalization with error propagation," *IEEE Trans. Commun.*, vol. 46, pp. 656–665, May 1998.
- [25] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Princeton, NJ: Van Nostrand, 1960.

- [26] J. Proakis, *Digital Communications*, 2nd ed: McGraw-Hill, 1989.
- [27] G. C. Clark Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*. New York: Plenum, 1981.
- [28] M. V. Eyuboğlu, "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interleaving," *IEEE Trans. Commun.*, vol. 36, pp. 401–409, Apr. 1988.
- [29] K. Zhou, J. G. Proakis, and F. Ling, "Decision-feedback equalization of time-dispersive channels with coded modulation," *IEEE Trans. Commun.*, vol. 38, pp. 18–24, Jan. 1990.

**Michael Reuter** (S'82–M'86) received the B.S. and M.S. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1984 and 1986, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, San Diego, in 2000.

Since 1987 he has been with the Space and Naval Warfare Systems Center, San Diego. His current research interests are in adaptive and statistical signal processing applied to problems in wireless communications.



**Jeffery C. Allen** (M'94) received the Ph.D. degree in mathematics from the University of California, San Diego, in 1988.

Since 1989 he has been a scientist with the Space and Naval Warfare Systems Center, San Diego. His current work is focused on  $H^\infty$  applications to electrical circuits, stochastic sea surfaces for RF propagation, and wireless channel modeling.



**James R. Zeidler** (M'76–SM'84–F'94) received the Ph.D. degree in physics from the University of Nebraska, Lincoln, in 1972.

Since 1974 he has been a scientist with the Space and Naval Warfare Systems Center, San Diego, CA. He has also been an Adjunct Professor of electrical and computer engineering with the University of California, San Diego, since 1988. During this period, he conducted research on communications systems, sonar and communications signal processing, communications signals exploitation,

undersea surveillance, array processing, underwater acoustic communications, infrared image processing, and electronic devices. He was also a technical advisor with the Office of the Assistant Secretary of the Navy (Research, Engineering and Systems), Washington, DC from 1983 to 1984.

Dr. Zeidler was an Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 1991 to 1994 and is a member of the Editorial Board of the *Journal of the Franklin Institute*. He received an award for the Best Unclassified Paper at the IEEE Military Communications Conference in 1995 and the Navy Meritorious Civilian Service Award in 1991.

**Richard C. North** (S'84–M'90) received the B.S. degree from the University of California, Davis, in 1984, the M.Eng. degree from Cornell University, Ithaca, NY, in 1985, and the Ph.D. degree from the University of California, San Diego, in 1991, all in electrical engineering.

Since 1992, he has been with the Space and Naval Warfare Systems Center, San Diego. From 1992 to 1994, he worked on the research and development of combined space-time adaptive filtering algorithms for advanced sonar and communication applications. From 1994 to 1999, he was the chief engineer on the High Data Rate Line-of-Sight maritime communications project. During this time, he was in charge of the development and demonstration of LOS communications equipment between ships and ship-to-shore at data rates up to several Mbps. Since 1998, he has been the technical lead on the Digital Modular Radio (DMR) program. DMR is a software programmable, multiband, multi-mode radio designed for surface and subsurface Naval platforms.