Evaluation of the Effects of Intersymbol Interference in Decision Feedback Equalizers

T. J. Willink, P. H. Wittke, Senior Member, IEEE,

L. L. Campbell, Life Fellow, IEEE *

January 21, 2000

*This research was supported by the Natural Sciences and Engineering Research Council of Canada, under grants OPG 0003391 and OPG 0002151. Some of these results were reported at the International Symposium on Information Theory in Trondheim, Norway, June 1994.

T. J. Willink is with the Communications Research Centre, 3701 Carling Ave., Box 11490, Station H, Ottawa, Ontario K2H 8S2.

P. H. Wittke is with the Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario K7L 3N6.

L. L. Campbell is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario K7L 3N6.

Abstract

For channels which suffer predominantly from additive noise and intersymbol interference, the decision feedback equalizer has provided a relatively simple solution for reducing the effects of interfering symbols at the input to the decision device. In this paper, a technique is developed which enables fast, accurate calculation of the error performance of decision feedback equalization for a number of channel models. The method is to calculate the n-step transition probability for an associated Markov process and then to use this transition probability as an approximation to the stationary probability distribution. For systems with finite memory it is proved that the method converges. If the signal-to-noise ratio is high and the signal amplitude is more than twice the worst-case interference, it is shown that the convergence is rapid. Numerical results indicate that the convergence is rapid enough to make this an efficient method of calculation even for channels for which the interference does not fully satisfy this condition. Two examples are given here, but the technique has been tested on most of the examples that have been presented in the literature. The method yields results in closer agreement with simulation results than previous results obtained using bounding techniques, especially at low to moderate signal-to-noise ratios, and requires less computation.

Index Terms: Decision feedback equalization, intersymbol interference.

1 Introduction

In channels with memory, intersymbol interference (ISI) limits system performance. Often, the intersymbol interference component in the decision statistic at time kT can be expressed as

$$Z_k = \sum_{n=1}^{\infty} \beta_n X_{k-n}, \tag{1}$$

where X_{k-n} are interference symbols which may be previous data symbols or some combination of data symbols and estimates of these symbols which come from a decision feedback equalizer (DFE), and $\{\beta_n\}$ is the sampled impulse response of the channel model. Even when the interference symbols are independent and identically distributed (i.i.d.), evaluating the effects of the ISI can be a substantial task.

Some progress has been made in evaluating the error probability in the presence of ISI when the interference symbols are independent and uniformly distributed. Hill and Blanco [1] presented a survey of results for the special case in which the ISI coefficients are exponentially decaying, i.e. $\beta_n = \beta^n$. If the symbols X_n take the values ± 1 with equal probability and $\beta = 1/2$, it is known that the distribution is uniform [2, sec.4.2]. Rice [3] evaluated the distribution for the case $\beta_n = 1/n$ using a numerical integration procedure.

A variety of bounds and approximations have been developed for the error probability in models with ISI caused by i.i.d. variables. They were summarized and discussed by Helstrom in [4]. A new method of evaluating the pdf of this type of ISI was also given by Metzger [5]. More recent work includes [6], in which the error probability for intersymbol and cochannel interference channels was evaluated using a Fourier series expansion of the noise distribution function and [7] where some further comments on these methods may be found. In [8] Wittke et al. gave a method for evaluating the distribution function, which was then used to calculate the error probability exactly. In [9], this procedure was improved by eliminating the need to calculate the distribution function explicitly, and by enlarging the class of models to which the technique applies. The method was further extended to apply to M-ary modulations in [10] and in [11] there are upper and lower bounds on the error probability as well as some further information about the distribution of Z_k.

Currently, little is known about the distribution function of the ISI when it is caused by random variables that are mutually dependent because they are the symbols in the feedback section of a DFE. (A number of authors, [12, 13] for example, have considered the bounding or calculation of error probability in the presence of ISI when the transmitter symbols are dependent because of correlative encoding, but this is a rather different problem from that considered here.) There is a considerable literature on the calculation of bounds on error probability in a DFE. The starting point for much of this literature is the paper of Duttweiler et al. [14]. For a recent bibliography of previous work on the calculation of error probability for the DFE, with emphasis on bounding techniques, see [15].

In this paper, the emphasis is on calculation of the error probability rather than on bounds. The method uses a well-known idea from Markov chains, which proves to be highly effective and computationally inexpensive in its application to the DFE. The idea is that the probability distribution in an ergodic process tends to a limiting stationary distribution as time goes on and that, for many processes, the approach to the stationary distribution is fairly rapid, especially if the initial distribution is chosen well. The method implicitly accounts for error propagation. A proof is given that the method is valid for channels with finite memory. It is shown in the Appendix that the method converges quickly for large signal-to-noise ratio (SNR) when the interference is small enough that no errors would occur in the absence of noise. It is a reasonable conjecture, for which there is good numerical evidence (see the examples and [16]), that the same method is quite accurate even for fairly low SNR and that it can be used for channels with infinite memory, as long as the series $\sum |\beta_n|$ converges absolutely.

The numerical procedure is systematic and provides significantly more accurate (as measured by agreement with simulations) results than those obtained previously by bounding techniques and with less computation. The improvement is particularly noticeable at lower values of SNR.

Examples are given in Sections 2.3 and 3.2 to demonstrate the method and to allow comparison with results obtained using bounding techniques. In the first example, there is finite memory and the DFE is matched to the channel. The second example is an example of a channel with infinite memory for which the DFE and channel are mismatched. Even though there is no proof of convergence given in this case, the calculated values agree with simulation results.

2 Ideal Decision Feedback Equalizer

2.1 Method of calculation

In this section we consider an "ideal" DFE model, i.e. one which eliminates all anti-causal interference and which cancels all other interference when no decision errors occur. Thus the only interference is due to decision errors. As in [15] and most of [14], it is assumed that a zero-forcing criterion has been used in the design of the DFE forward filter.(In Sec. 3, we extend this to the case where the feedback filter does not exactly match the channel characteristics so that interference occurs even when all decisions are correct.) We also assume that the channel and DFE have finite memory M, so that the series in (1) has only M non-zero terms. For simplicity, we consider only binary transmission of data variables that are independent of one another. An extension to other than binary transmission presents no difficulty in principle. Our objective is to obtain an expression for the probability of error in decisions when the DFE is employed.

Consider the DFE block diagram of Fig. 1, where h_i represents the sampled impulse response of the cascade of the transmit filter, the physical medium, receive filter, and the DFE forward filter. Because the feedback filter is assumed to be ideal, $\beta_i = h_i$ for $i = 1, 2, \ldots, M$. Let the signal amplitude be $h_0 = A$. Let the input data sequence be $\{\ldots, a_{-2}, a_{-1}, a_0, a_1, \ldots\}$, where the data variables a_k are i.i.d. random variables with $P\{a_k = +1\} = P\{a_k = -1\} = 1/2$. The input to the decision device at time k is

$$\tilde{a}_k = Aa_k + Z_k + N_k, \tag{2}$$

where the noise variables N_k are assumed to be independent Gaussian random variables with mean 0 and variance $\sigma^2 > 0$ and the ISI variables Z_k are caused by previous decision errors. The decision device makes the decision $\hat{a}_k = +1$ if $\tilde{a}_k$ is positive and makes the decision $\hat{a}_k = -1$ otherwise. This model is similar to that studied by Duttweiler et al. [14] and by Kennedy et al. [17] although here the noise is taken to be Gaussian and our objectives are rather different from those of [17]. Define the error variable $e_k = a_k - \hat{a}_k$, so that e_k takes values in the set $\{-2, 0, +2\}$. Note that the random variables e_{k-i} are dependent because each error depends on previous errors. Moreover, they are not uniformly distributed, since, except at very low SNR, the event $e_k = 0$ should be more probable than $e_k = \pm 2$. Thus most previous theoretical results on the distribution of ISI are not applicable to this situation.

Because of the finite memory, the state at time k can be described by the vector

$$\mathbf{f}_k = (e_{k-1}, e_{k-2}, \ldots, e_{k-M}).$$

It can be seen that the states $\{\mathbf{f}_k\}$ form a Markov chain with transition probabilities which depend on the ISI variable

$$Z_k = Z(\mathbf{f}_k) = \sum_{i=1}^{M} \beta_i e_{k-i}. \tag{3}$$

While $\mathbf{f}_k$ can have 3^M values, $\mathbf{f}_{k+1}$ can have only three values for any given $\mathbf{f}_k$, namely the vectors

$$(e_k, e_{k-1}, \ldots, e_{k-M+1})$$

for $e_k = -2, 0, 2$. We denote these vectors $[-2, \mathbf{f}_k], [0, \mathbf{f}_k]$, and$[2, \mathbf{f}_k]$ respectively. Note that

$[-2, \mathbf{f}_k]$ etc. are M-tuples formed by dropping the entry e_{k-M} in $\mathbf{f}_k$, shifting the remaining components one place to the right, and adding a new first component.

The transition probabilities of the Markov chain are easily found. We have

$$P\{\mathbf{f}_{k+1} = [-2, \mathbf{f}_k] \big| \mathbf{f}_k\} = P\{a_k = -1, \hat{a}_k = 1 \big| \mathbf{f}_k\} = P\{a_k = -1\} P\{\hat{a}_k = 1 \big| a_k = -1, \mathbf{f}_k\}$$
$$= \frac{1}{2} P\{-A + Z(\mathbf{f}_k) + N_k > 0\} = \frac{1}{2} \Phi\left(\frac{-A + Z(\mathbf{f}_k)}{\sigma}\right), \quad (4)$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Similarly,

$$P\{\mathbf{f}_{k+1} = [2, \mathbf{f}_k] \big| \mathbf{f}_k\} = \frac{1}{2} \Phi\left(\frac{-A - Z(\mathbf{f}_k)}{\sigma}\right) \quad (5)$$

and

$$P\{\mathbf{f}_{k+1} = [0, \mathbf{f}_k] \big| \mathbf{f}_k\} = 1 - \frac{1}{2}\left[\Phi\left(\frac{-A + Z(\mathbf{f}_k)}{\sigma}\right) + \Phi\left(\frac{-A - Z(\mathbf{f}_k)}{\sigma}\right)\right]. \quad (6)$$

The other transition probabilities are zero.

It is possible to go from any state $\mathbf{E}_a = (a_1, \ldots, a_M)$ to any other state $\mathbf{E}_b = (b_1, \ldots, b_M)$ in a finite number number of steps with positive probability. (One way to do this is to go from $\mathbf{E}_a$ to $\mathbf{0} = (0, \ldots, 0)$ and then from there to $\mathbf{E}_b$, although this may not be the only path.) The Markov chain is thus irreducible in the usual terminology [18]. The chain is also aperiodic. (If state $\mathbf{E}_a$ had period t then returns to the state would be possible only at time intervals $t, 2t, \ldots$. However the system can go from $\mathbf{E}_a$ to $\mathbf{0}$, remain there an arbitrary length of time and then return to $\mathbf{E}_a$, showing that there is no period.) For any finite, irreducible, aperiodic Markov chain, there is a stationary distribution, independent of the initial state, to which the n-step transition probabilities tend as $n \to \infty$ [18, p. 357].

To calculate the steady-state probability of error one could find the stationary distribution as the solution q of the matrix equation

$$qP = q \tag{7}$$

where P is the $3^M \times 3^M$ transition matrix of the chain. Although this calculation can be simplified some by using the symmetry inherent in the model [19], for large values of M this computation is unattractive. For moderate values of M, solution of (7) might be facilitated by use of the ideas in the review paper [20] by Mayer. This direction is not pursued here.

The method which is used in this paper is as follows:

(i) Consider a system which begins at time zero in the state $\mathbf{f}_0 = \mathbf{0}$. This choice of initial state is not essential, but it should speed convergence for the high SNR case, because the most probable states are the zero state and those close to it.

(ii) Use standard Markov chain results to calculate the n-step transition probability

$$P^{(n)}\{\mathbf{f}_n = \mathbf{e} \big| \mathbf{f}_0 = \mathbf{0}\} = \sum_{\mathbf{f}_{n-1},\ldots,\mathbf{f}_2,\mathbf{f}_1} P\{\mathbf{f}_n = \mathbf{e} \big| \mathbf{f}_{n-1}\} P\{\mathbf{f}_{n-1} \big| \mathbf{f}_{n-2}\} \cdots P\{\mathbf{f}_1 \big| \mathbf{f}_0 = \mathbf{0}\} \tag{8}$$

where the conditional probabilities are given by (4) – (6) and the summation is over all the values of $(\mathbf{f}_{n-1}, \ldots, \mathbf{f}_2, \mathbf{f}_1)$ that are possible with the initial value $\mathbf{f}_0 = 0$ and the final value $\mathbf{f}_n = \mathbf{e}$.

(iii) Treat the n-step transition probability from step (ii) as if it were the stationary probability and calculate the probability of bit error at time n, i.e. the probability that $\mathbf{f}_{n+1} = [-2, \mathbf{e}]$ or $[2, \mathbf{e}]$ for some $\mathbf{e}$,

$$P\{\text{error}\} = \frac{1}{2} \sum_{\mathbf{e}} P^{(n)}\{\mathbf{f}_n = \mathbf{e} \big| \mathbf{f}_0 = \mathbf{0}\} \left[\Phi\left(\frac{-A + Z(\mathbf{e})}{\sigma} \right) + \Phi\left(\frac{-A - Z(\mathbf{e})}{\sigma} \right) \right]. \qquad (9)$$

Because of the convergence of transition probabilities for finite irreducible aperiodic Markov chains, for large enough n this will be close to the error probability which would be calculated from the stationary distribution.

If the SNR is large and $|Z(\mathbf{f}_k)| < A$ for all $\mathbf{f}_k$, then, as is shown in the Appendix, the convergence to the stationary distribution is rapid as n increases. The condition $|Z(\mathbf{f}_k)| < A$ implies that no errors will occur in the absence of noise. This condition is comparable to an eye-open condition at the decision threshold. But it should be noted that $Z(\mathbf{f}_k)$ in (3) depends on the decision errors instead of the input interference alone.

Convergence occurs whether or not $|Z(\mathbf{f}_k)| < A$ for all $\mathbf{f}_k$; it is only the speed of convergence that is discussed in the Appendix. The result in the Appendix suggests, but does not prove, that the convergence should still be moderately rapid if the condition $|Z(\mathbf{f}_k)| < A$ is violated for some less probable error patterns $\mathbf{f}_k$. The calculations for Example I, Sec. 2.3, confirm that this is so, at least for this example.

Note that this method can be used for arbitrarily large memory M. If $M = \infty$, the same calculations can be performed, because for $\mathbf{f}_0 = \mathbf{0}$, the sums in (8) and (9) are finite sums for each finite n. However, for $M = \infty$, it is *not* proved here that $P^{(n)}$ converges to a stationary distribution as $n \to \infty$, since the theorem used above applies only to finite state Markov chains. The number of states for $M = \infty$ can be uncountably infinite, since, as follows by arguments similar to those in [8], the set of states can be Cantor-like. It is conjectured that convergence occurs under conditions likely to occur in practice, but a proof is not known to the authors.

2.2 Computational issues

For the two examples which are given in this paper, and all others in [16], the value $n = 4$ (and sometimes less) gives very good agreement with simulations, even at quite moderate values of signal-to-noise ratio. The examples in [16] include the example in [14] and all three examples used by Altekar and Beaulieu [15] , although, to conserve space, only one of the three examples is presented here.

In step (ii) above, $\mathbf{e}$ is any allowable state vector at time n, given that the initial state is $\mathbf{0}$ and the sum is over all possible $(\mathbf{f}_{n-1}, \ldots, \mathbf{f}_2, \mathbf{f}_1)$ which lead to $\mathbf{f}_n = \mathbf{e}$ from $\mathbf{f}_0 = \mathbf{0}$. When $n \leq M$, there is only one such n-tuple of vectors $\mathbf{f}_i$ and the sum (8) consists of just one term. This term is a product of n factors. The sum in (9) is over allowable $\mathbf{e}$, i.e. those $\mathbf{e}$ which can be reached in n steps after starting with $\mathbf{f}_0 = 0$. Thus, for $n \leq M$, there are 3^n terms in this sum and, for $n > M$ there are 3^M terms. To put these numbers in some perspective, in our examples $n = 4$ gives good results for memories M of 10 and ∞. Hence the calculation of (9) involved summing 81 terms, while solving (7) even with the symmetry taken into account as in [19] would involve a matrix with $(3^{10} + 1)/2 = 29525$ rows and columns.

2.3 Example I

To allow a comparison with results obtained using previous methods, the first example is taken from [15], where it is called 'Channel B'. The signal is binary PAM, with data symbols $a_k \in \{-1, +1\}$, the DFE is ideal and the finite channel response is given by

$$\beta_i = h_i = \begin{cases} \frac{1}{i+1} & 0 \leq i \leq 10 \\ 0 & \text{otherwise.} \end{cases}$$

To examine the accuracy of the theoretical calculations, the system was simulated using a Monte-Carlo method. The simulations were run until the cumulative error rate remained within a range of 1 % over five successive blocks of 10,000 symbols. In Fig. 2, the calculated error probability is shown for $n = 4$ along with the simulated performance, where n is the point at which it was assumed that $P^{(n)}\{\mathbf{f}_n \mid \mathbf{f}_0 = \mathbf{0}\}$ is approximately equal to the steady-state distribution. The calculated curve appears coincident with the simulated error rates. As well, the ISI-free lower bound and the best upper bound of [15], the Altekar-Beaulieu (A-B) bound, are shown. The computational effort required to calculate the A-B bound was significantly greater than that required to calculate the error probability from (9). As was demonstrated in [15], the A-B bound is the tightest known for this channel. While it is reasonably tight at moderate to large SNR, it is poor at low SNR.

Error probabilities were calculated for the example, using (9) with n=2, 3, 4, and 5. The error probability curves converge fairly rapidly as n is increased. The maximum deviation from the simulated results is approximately 0.2 dB for the calculation with $n = 2$.

For $n = 4$ and $n = 5$, the calculated error rates are very close to the simulated performance for the SNR range 0–10 dB and the error-rate curves appear indistinguishable. For more details, see [16]. In [16], the technique presented in this paper was applied to the other two examples in [15]. In both cases, the calculated error probability very closely approximated the simulated performance for the full range of signal-to-noise ratios and gave a significantly more accurate result than the A-B upper bound.

3 Non-Ideal Decision Feedback Equalizer

3.1 Outline of method

In many cases of practical interest, there may be uncancelled ISI. This residual ISI will arise when the feedback filter has insufficient taps to cancel all the interfering terms, or when the filter is not matched exactly to the channel, for example in DFE's when the channel is time-varying or the identification is inaccurate. When residual ISI remains, it is not independent of the interference due to the feedback sequence $\{e_k\}$.

In this section, an extension is outlined to the theory presented so far to enable the effects of dependent, causal, residual ISI to be taken into account. The development in this section again assumes a one-dimensional signalling constellation, i.e. binary PAM.

If the feedback filter is mismatched to the channel, i.e. not all the causal ISI is eliminated, the feedback filter coefficients can be written

$$\beta_i = h_i + \Delta h_i \quad \text{if } i \geq 1,$$

where Δh_i is the error due to the time-varying channel or identification inaccuracies. If the feedforward filter removes all the anti-causal ISI, the DFE equation becomes

$$\tilde{a}_k = Aa_k + \sum_{i=1}^{\infty} h_i a_{k-i} - \sum_{i=1}^{\infty} \beta_i \hat{a}_{k-i} + N_k. \tag{10}$$

The error variable is defined as before, i.e. $e_k = a_k - \hat{a}_k$. The interference variable is then

$$Z_k = \sum_{i=1}^{\infty} \beta_i e_{k-i} - \sum_{i=1}^{\infty} \Delta h_i a_{k-i}. \tag{11}$$

This will be considered as two separate summations,

$$Z_k^{(1)} = \sum_{i=1}^{\infty} \beta_i e_{k-i} \quad \text{and} \quad Z_k^{(2)} = \sum_{i=1}^{\infty} \Delta h_i a_{k-i}. \tag{12}$$

We remark that if we replaced a_{k-i} by a_{k+i} and Δh_i by $-h_{-i}$ in $Z_k^{(2)}$ in (12) we would have a model for including the ISI that would be present if the forward filter were not zero-forcing.

While the details of the calculation are more complicated, the same approach as in the previous section can be used. We assume that the process starts at time zero with all interference variables set equal to zero. The process is then allowed to evolve to time n with the step-by-step probabilities calculated in a manner similar to the previous section and finally the error probability is calculated. If n is large enough, it is assumed that the calculated probability closely approximates the steady-state probability of error.

3.2 Example II

An example is now presented in which the feedback filter of the DFE does not cancel all of the causal ISI due to identification errors. It is assumed that the feedforward filter does cancel all anti-causal ISI, and both the channel model and the feedback filter have infinite impulse responses.

The channel model is the exponential decay model used in [1] and [14, Sec. IV]

$$h_i = e^{-i\alpha} \qquad \text{if } i \geq 0, \tag{13}$$

where in this particular example we have set $\alpha = 0.9$ and the additive noise is white and Gaussian.

The coefficients of the feedback filter are

$$\beta_i = h_i(1+\delta) \qquad i \geq 1$$

where δ represents the inaccuracy of the feedback filter. (It is recognized that this model oversimplifies on reality, but it seems useful for gaining insight into the behavior of a DFE.) As in (11) and (12), the interference is $Z_k^{(1)} - Z_k^{(2)}$ where

$$Z_k^{(1)} = \sum_{i=1}^{\infty} \beta_i e_{k-i} \qquad Z_k^{(2)} = \delta \sum_{i=1}^{\infty} h_i a_{k-i}.$$

For binary PAM, the error probability conditioned on the ISI is

$$P\{e_k \neq 0 \Big| Z_k^{(1)}, Z_k^{(2)}\} = \frac{1}{2} P\{N_k \geq A + Z_k^{(1)} - Z_k^{(2)}\} + \frac{1}{2} P\{N_k > A - Z_k^{(1)} + Z_k^{(2)}\}. \tag{14}$$

The results of the error probability calculations and simulations are shown in Fig. 3 for three values of δ. The calculations were performed for $n = 4$ time steps. The value $\delta = 0$ corresponds to the case in which the feedback filter is exactly matched to the channel. The results for two mismatched DFE's are also shown, in which the feedback filter coefficients are 30% and 50% greater than the model coefficients (i.e. $\delta = 0.3,\ 0.5$). These curves give the performance of the DFE when the identification procedure is known to yield errors of 30% and 50% respectively. For the simulations, the infinite-length impulse response was truncated at a point such that the effect of the omitted terms was negligible.

It is seen that the calculated error probabilities agree very closely with the simulated error rates. The model with 30% identification errors is approximately 0.5 dB worse than the ideal DFE at an error probability of 10^{-2}, and the model with 50% identification errors is about 1.3 dB worse than the ideal DFE at the same error rate.

4 Conclusion

A new systematic approach for calculating the error probability of a decision feedback equalizer has been presented. The method is to calculate n-step transition probabilities for a Markov process associated with the errors in the decision element. The transition probabilities are used as an approximation to the stationary probability distribution of the Markov process. Error probability is then calculated from the approximate stationary distribution. It appears that good results are obtained for quite small values of n and thus quite modest amounts of calculation. The analysis shows that the method converges for finite-memory channels at all SNR and that the convergence is rapid if the SNR is large and the interference is small. The numerical evidence indicates that some of the mathematical restrictions (finite memory, very small interference, large SNR) can be relaxed without losing the good convergence properties.

The technique is applicable to a wide class of channels, and is readily extended to systems with more than one ISI variable, such as non-ideal DFE's or multi-dimensional signalling constellations. The technique has been applied to a variety of channel models, and the examples presented here have demonstrated the speed of convergence and the accuracy of the error probability estimates.

Acknowledgment

The authors are grateful to the referees and editor for pointing out some errors and for a number of suggestions to clarify the presentation.

Appendix. Speed of Convergence to Stationary Distribution

For the model of Sec.2.1, if $|Z(\mathbf{f}_k)| < A$ for all possible $\mathbf{f}_k$, i.e. if

$$2\sum_{i=1}^{M} |\beta_k| < A,$$

then

$$\lim_{\sigma \to 0^+} \Phi\left(\frac{-A \pm Z(\mathbf{f}_k)}{\sigma}\right) = 0.$$

Under this condition, the limiting values of the transition probabilities to $[0, \mathbf{f}_k]$ in (6) are

$$P\{\mathbf{f}_{k+1} = [0, \mathbf{f}_k] \Big| \mathbf{f}_k\} = 1$$

and the limiting values of the other transition probabilities are zero.

Denote by $P(\sigma)$ the matrix of transition probabilities given by (4)–(6) and by $P(0)$ the limiting value as $\sigma \to 0^+$. Similarly, denote by $\lambda_i(\sigma)$ and $\lambda_i(0)$ the eigenvalues of these matrices. It is known [21, p.371] that the eigenvalues of a matrix are continuous functions of the elements of the matrix. Hence $\lambda_i(\sigma)$ is a continuous function of σ which approaches $\lambda_i(0)$.

We now adopt an ordering of the 3^M states in which $\mathbf{0}$ is the first state, the next two are $(0, \ldots, 0, -2)$ and $(0, \ldots, 0, 2)$, the next six have $M-2$ leading zeros, the next eighteen have $M-3$ leading zeros and so on. Then, when $\sigma = 0$, the only possible transitions are to lower numbered states except for the first state, which has a transition to itself. All of the allowable transitions have probability 1, so that the transition matrix $P(0)$ has a single 1 on the diagonal in the $(1,1)$ position, with zeros everywhere else on and above the diagonal. Consequently the characteristic polynomial of $P(0)$ is

$$\det[\lambda I - P(0)] = \lambda^{3^M - 1}(\lambda - 1).$$

Thus, $P(0)$ has a simple eigenvalue equal to 1 and all its other eigenvalues are equal to 0.

Because of the continuity of eigenvalues as functions of σ this result about $P(0)$ implies that $P(\sigma)$ has $3^M - 1$ eigenvalues which are close to 0 for σ close to 0, i.e. for large SNR. The other eigenvalue must be 1, since 1 is always an eigenvalue of a transition matrix.

Without significant loss of generality, we assume that the eigenvalues of $P(\sigma)$ are distinct for $\sigma > 0$. Let the corresponding left and right eigenvectors be $\phi_1, \phi_2, \ldots$ and $\theta_1, \theta_2, \ldots$, normalized so that $\phi_i \theta_j = \delta_{ij}$. Note that, with our numbering, $\lambda_1(\sigma) = 1$ and so ϕ_1 is the stationary probability distribution and $\theta_1 = (1, 1, \ldots, 1)^T$. Any initial distribution $\mathbf{v}$ can be written as a row vector

$$\mathbf{v} = \sum_{k=1}^{3^M} c_k \phi_k, \tag{15}$$

where $c_k = \mathbf{v}\theta_k$. In particular, $c_1 = 1$ because $\mathbf{v}$ is a probability.

After n steps the transition probabilities, as in (8), are the elements of $P^n(\sigma)$. Recall that the eigenvalues of $P^n(\sigma)$ are $(\lambda_k(\sigma))^n$, that $\lambda_1(\sigma) = c_1 = 1$, and multiply (15) on the right by $P^n(\sigma)$ to get the probability after n steps

$$\mathbf{v}P^n(\sigma) = \phi_1 + \sum_{k=2}^{3^M} c_k (\lambda_k(\sigma))^n \phi_k \approx \phi_1,$$

where the last approximation follows because the other terms, with eigenvalues close to zero, approach zero exponentially fast with increasing n. Thus, no matter what the initial distribution, the probabilities converge quickly to the stationary distribution when the signal-to-noise ratio is large and $|Z(\mathbf{f}_k)| < A$ for all $\mathbf{f}_k$.

References

[1] F.S. Hill and M.A. Blanco, "Random geometric series and intersymbol interference", *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 326–335, May 1973.

[2] R.W. Lucky, J. Salz and E.J. Weldon Jr., *Principles of Data Communication*, McGraw-Hill, 1968.

[3] S.O. Rice, "Distribution of $\sum a_n/n$, a_n randomly equal to ± 1", *Bell Syst. Tech. J.*, vol. 52, pp. 1097–1103, Sep. 1973.

[4] C.W. Helstrom, "Calculating error probabilities for intersymbol and cochannel interference", *IEEE Trans. Commun.*, vol. 34, pp. 430–435, May 1986.

[5] K. Metzger, "On the probability density of intersymbol interference", *IEEE Trans. Commun.*, vol. 35, pp. 396–402, Apr. 1987.

[6] N.C. Beaulieu, "The evaluation of error probabilities for intersymbol and cochannel interference", *IEEE Trans. Commun.*, vol. 39, pp. 1740–1749, Dec. 1991.

[7] L.L. Campbell and P.H. Wittke, "Mathematical problems in error calculations for interference", *IEEE Trans. Commun.*, vol. 45, pp. 1527–1528, Dec. 1997.

[8] P.H. Wittke, W.S. Smith and L.L. Campbell, "Infinite series of interference variables with Cantor-type distributions", *IEEE Trans. Inform. Theory*, vol. 34, pp. 1428–1436, Nov. 1988.

[9] W.S. Smith, P.H. Wittke and L.L. Campbell, "Error probabilities on fading channels with intersymbol interference", *IEEE Trans. Inform. Theory*, vol. 39, pp. 1598–1607, Sep. 1993.

[10] L.L. Campbell, P.H. Wittke and A.L. McKellips, "Intersymbol interference calculations for multi-level amplitude and phase modulated signals", URSI International Symposium on Signals, Systems and Electronics, Paris, 1992, pp.441–445.

[11] L.L. Campbell, A.L. McKellips and P.H. Wittke, "Distributions and expectations of singular random variables", *J. Appl. Prob.*, vol. 32, pp. 1032–1040, Dec. 1995.

[12] I. Korn, "Bounds to probability of error in binary communications systems with intersymbol interference and dependent or independent symbols," *IEEE Trans. Commun.*, vol. COM-22, pp. 251–254, Feb. 1974.

[13] G. L. Cariolaro and S.G. Pupolin, "Considerations on error probability in correlated-symbol systems," *IEEE Trans. Commun.*, vol. COM-25, pp. 462–467, April 1977.

[14] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmidt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 490–497, July 1974.

[15] 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.

[16] T.J. Willink, "Multicarrier transmission and intersymbol interference in decision feedback equalisers", Ph.D. thesis, Queen's Univ., Kingston, Ont., Nov. 1993.

[17] R. A. Kennedy, B.D.O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers", *IEEE Trans. Commun.*, vol. COM-35, pp.1022–1028, Oct. 1987.

[18] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 2nd ed. Wiley, 1961.

[19] 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.

[20] C. D. Mayer, "Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems," *SIAM Review*, vol. 31, pp. 240–272, June 1989.

[21] P.Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, 1985.

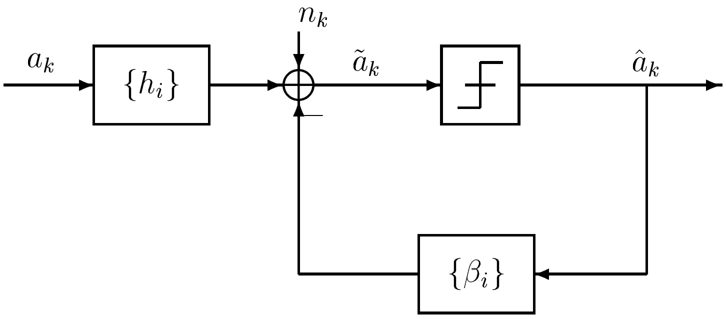

Figure 1: Decision Feedback Equalizer

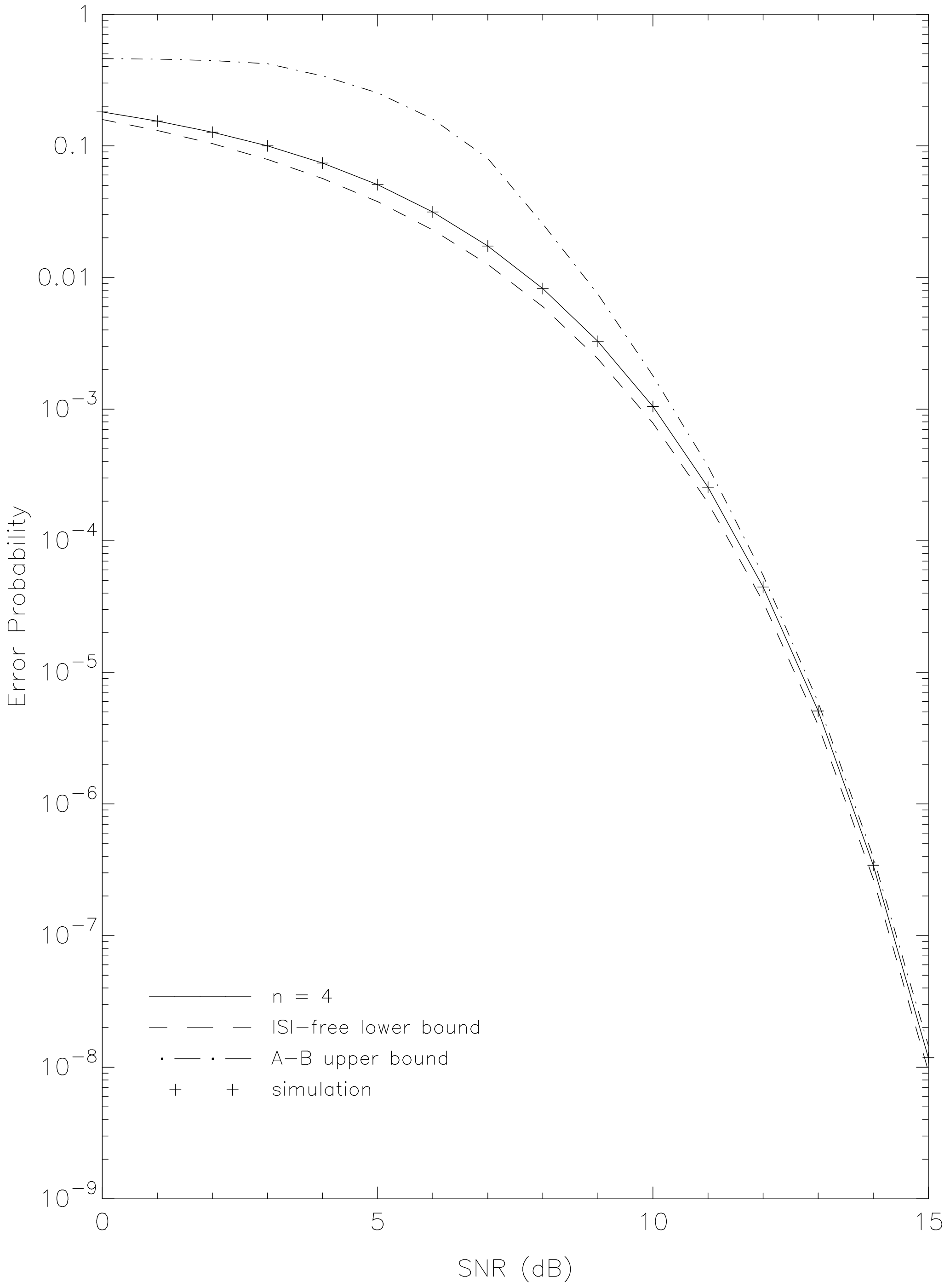

Figure 2: **Probability of Error for Example I**

Probability of error evaluated for $n = 4$. The simulated performance, the ISI-free lower bound and the best upper bound of [15] are also shown.

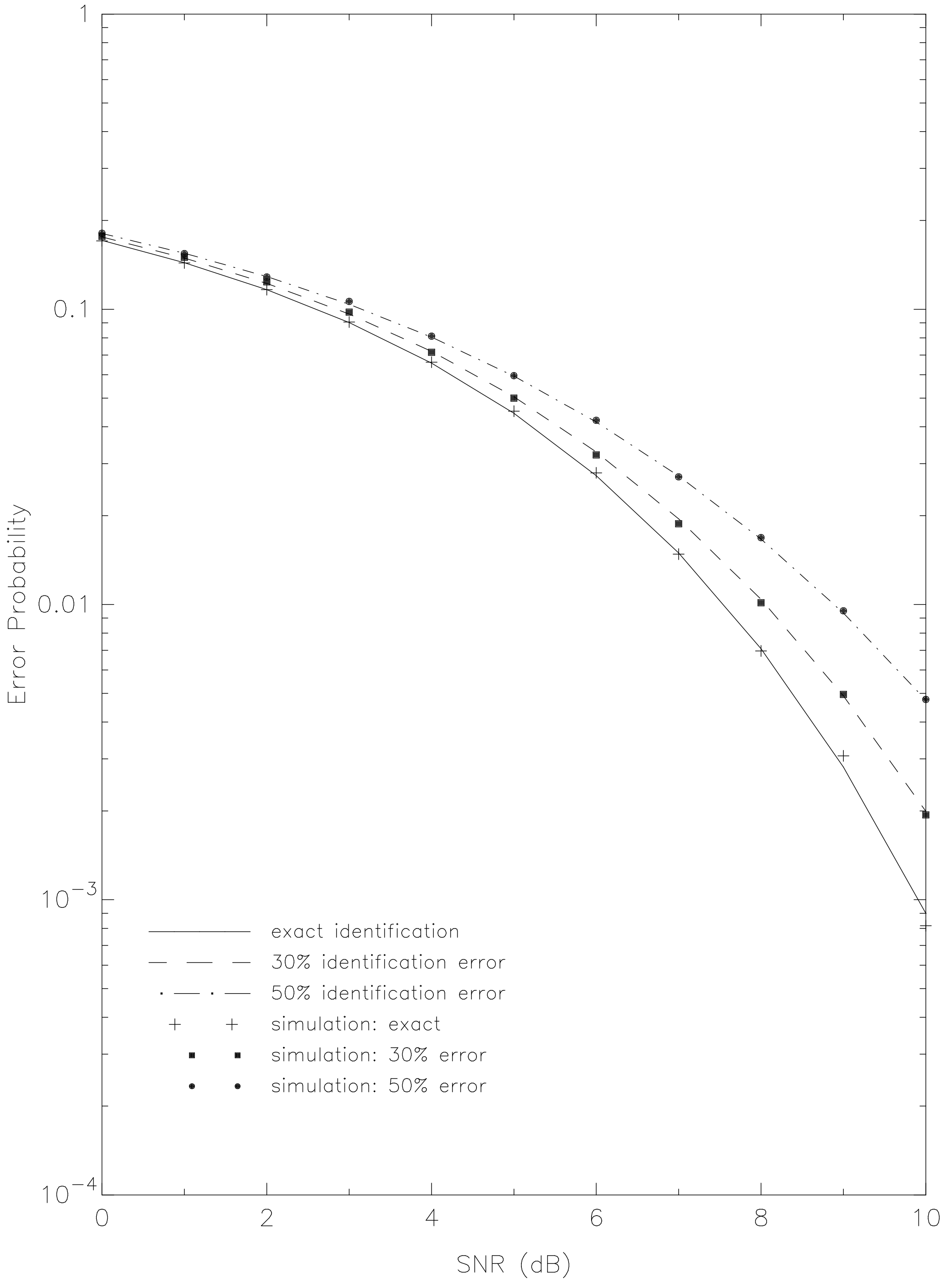

Figure 3: **Probability of Error for Example II** Probability of error evaluated for $n = 4$ and $\delta = 0,\ 0.3,\ 0.5$. The simulated performance is also shown.