Performance Analysis of Turbo-Equalized Partial Response Channels

Mats Öberg, *Member, IEEE,* and Paul H. Siegel, *Fellow, IEEE*

***Abstract*—The performance of maximum-likelihood decoding of a serial concatenation comprising a high-rate block code, convolutional code, or a turbo code, a uniform interleaver, and a partial response channel with additive white Gaussian noise will be addressed. The effect of a channel precoder on the system performance is also considered. Bit- and word-error rate estimates based upon properties of the average Euclidean distance spectrum of the coded partial response channel are derived. The estimates are compared to computer simulation results, and implications for system design are discussed.**

***Index Terms*—Digital magnetic recording, partial response channels, turbo-equalization.**

I. Introduction

TRELLIS-CODING techniques that improve the reliability of binary input-constrained, intersymbol interference (ISI) channels are of interest in both digital communications and data storage applications. Drawing inspiration from the success of turbo codes [1], [2], several authors have recently considered iterative decoding architectures for coding schemes of the form depicted in Fig. 1, where the outer encoder is a block, convolutional, or turbo encoder, π is an interleaver, $g(D)$ represents a precoder function, and $h(D)$ is the channel transfer polynomial.

This system resembles serial concatenation of interleaved codes, investigated by Benedetto *et al.* [3], with the inner code replaced by the ISI channel. For such a system, Douillard *et al.* [4] presented an iterative receiver structure, dubbed "turbo-equalization," to combat ISI due to multipath effects on convolutionally coded Gaussian and Rayleigh transmission channels. They introduced an interleaver between the encoder and channel, and as in turbo decoding, soft-output decisions from the channel detector and from the convolutional decoder were used in an iterative and cooperative fashion to generate estimates of the transmitted data.

Paper approved by W. E. Ryan, the Editor for Modulation, Coding, and Equalization of the IEEE Communications Society. Manuscript received July 19, 1999; revised April 3, 2000 and July 12, 2000. This work was supported in part by the National Science Foundation under Grant NCR-9612802, by UC MICRO Grant 98-140 in cooperation with Marvell Semiconductor, Inc., by the Center for Magnetic Recording Research at UCSD, by the Sweden-America Foundation, and by the Foundation Blanceflor Boncompagni-Ludovisi, née Bildt. This paper was presented in part at the 36th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 1998.

M. Öberg was with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA. He is now with Marvell Semiconductor, Inc., Sunnyvale, CA 94085 USA (e-mail: moberg@marvell.com).

P. H. Siegel is with the Center for Magnetic Recording Research, University of California, San Diego, La Jolla, CA 92093-0401 USA (e-mail: psiegel@ucsd.edu).

Publisher Item Identifier S 0090-6778(01)02172-9.

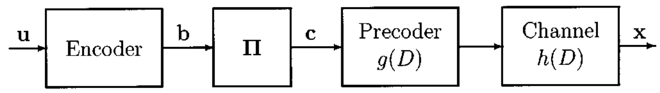

Fig. 1. Trellis-coded partial response system.

Motivated largely by the potential applications to digital magnetic recording, several authors have explored turbo-coding methods for the dicode and class IV partial response (PR4) channels, which have transfer functions $h(D) = 1 - D$ and $h(D) = 1 - D^2$, respectively. Heegard [5] and Pusch *et al.* [6] illustrated the design and iterative decoding of turbo codes for the dicode channel, using rates 1/2 and lower. Reed and Schlegel [7], extending prior results on a low-complexity, iterative multiuser receiver structure with interference cancellation, have evaluated the benefits of turbo-equalization for a rate 1/2, convolutionally coded, PR4 channel and E^2PR4 channel.

Ryan *et al.* [8], and Ryan [9], demonstrated that by using as an outer code a parallel-concatenated turbo code, punctured to achieve rates 4/5, 8/9, and 16/17 typical of commercial magnetic recording systems, one could obtain significant coding gain relative to previously known high-rate trellis-coding techniques on a precoded dicode or PR4 channel.

Recently, Souvignier *et al.* [10] and McPheters *et al.* [11] considered serial concatenated systems similar to that addressed in [8] and [9]. They investigated, by means of computer simulation, the performance achievable on a precoded dicode channel, with a high-rate convolutional code, rather than a turbo code, as the outer code. Somewhat surprisingly, the convolutional code was found to perform as well as the turbo code. Moreover, removal of the channel precoder was found to improve the performance of the turbo-coded system at low signal-to-noise ratio (SNR), while degrading the performance of the convolutionally coded system.

This paper was motivated, in part, by the desire to better understand the empirically observed differences in error rate in the precoded and nonprecoded serial concatenated systems. We address the performance of maximum-likelihood (ML) decoding of a serial concatenated system as shown in Fig. 1. The system comprises a high-rate block code, convolutional encoder, or parallel concatenation of two convolutional encoders as the outer code, an interleaver, and a partial response channel, with and without a channel precoder. The channel output is corrupted by additive white Gaussian noise (AWGN).

Let m denote a codeword, and let $\mathbf{x}_m$ denote the corresponding noiseless channel output. The Euclidean distance between the noiseless channel output sequences corresponding to codewords m and m' is $d_E = ||\mathbf{x}_m - \mathbf{x}_{m'}||$.

The ML union bound on word-error rate (WER) for a block-code on an AWGN channel with mean zero and variance σ^2, where all codewords are equally likely, is given by [12]

$$P_E \leq \frac{1}{M}\sum_{m=1}^{M}\sum_{m'\neq m} Q\left(\frac{\|\mathbf{x}_m - \mathbf{x}_{m'}\|}{2\sigma}\right) \tag{1}$$

where M denotes the number of codewords. We let $T_{m,\,d_E}$ denote the number of noiseless channel output sequences $\mathbf{x}_{m'}$ that lie at Euclidean distance d_E from $\mathbf{x}_m$. Then, we can write (1) as follows:

$$\begin{aligned} P_E &\leq \frac{1}{M}\sum_{m=1}^{M}\sum_{d_E=1}^{\infty} T_{m,\,d_E} Q\left(\frac{d_E}{2\sigma}\right) \\ &= \sum_{d_E=1}^{\infty}\frac{1}{M}\sum_{m=1}^{M} T_{m,\,d_E} Q\left(\frac{d_E}{2\sigma}\right) \\ &= \sum_{d_E=d_{\min}}^{\infty} \overline{T}(d_E) Q\left(\frac{d_E}{2\sigma}\right) \end{aligned} \tag{2}$$

where $\overline{T}(d_E) = (1/M)\sum_{m=1}^{M} T_{m,\,d_E}$ is the average number of noiseless channel output sequences $\mathbf{x}_{m'}$ at Euclidean distance d_E from a given sequence $\mathbf{x}_m$. The corresponding bit-error rate (BER) bound can be derived similarly

$$P_b \leq \sum_{d_E=d_{\min}}^{\infty} \frac{\overline{T}(d_E)\overline{w}(d_E)}{K} Q\left(\frac{d_E}{2\sigma}\right) \tag{3}$$

where K denotes the number of information bits represented by a codeword m, and $\overline{w}(d_E)$ denotes the average Hamming distance between information words that generate codewords m and m' whose corresponding noiseless channel output sequences $\mathbf{x}_m$ and $\mathbf{x}_{m'}$ lie at Euclidean distance d_E.

For an exact analysis, we must determine the full compound error-event characterization for a code interleaved and concatenated with the ISI channel. The complexity of this computation is generally prohibitively high. To overcome this difficulty, we introduce a technique for computing an approximation to the average weight enumerator $\overline{T}(d_E)$ for a high-rate, coded partial response-channel. The result depends only upon the output Hamming weight enumerator of the outer code

$$A(d) = \sum_{i=0}^{K} A(d,\, i) \tag{4}$$

where $A(d,\, i)$ denotes the number of error words of Hamming output weight d and input weight i.

In Section II, we present the error-event analysis for the dicode channel, first without a precoder, then with a precoder of the form $g(D) = 1/(1 \oplus D)$. We also extend the analysis to the PR4 channel. In Section III, we consider dicode systems incorporating a rate 8/9 outer punctured convolutional code and a rate 4/5 turbo code. The performance estimates based upon the analysis in Section II are compared to the results of computer simulation. Section IV concludes the paper.

II. Error Event Analysis on the Dicode Channel

Referring to the system model in Fig. 1, we assume that the encoder is a block encoder, for example, a truncated convolutional encoder or a turbo encoder. Let $\mathbf{b} = [b_1, \ldots, b_N]$ denote a codeword, and $\mathbf{c} = \pi(\mathbf{b})$ be the corresponding output of the interleaver. The output of the channel is denoted $\mathbf{x} = [x_1, \ldots, x_N]$. Given two codewords $\mathbf{b_1}$ and $\mathbf{b_2}$, let $\mathbf{e} = \mathbf{b_1} \oplus \mathbf{b_2}$ be the corresponding Hamming error word, and let $\mathbf{f} = \mathbf{c_1} \oplus \mathbf{c_2}$ be the interleaved Hamming error word. Let $\boldsymbol{\epsilon} = \mathbf{b_1} - \mathbf{b_2}$ be the signed error word, with corresponding interleaved signed error word $\boldsymbol{\zeta} = \mathbf{c_1} - \mathbf{c_2}$ and channel output error word $\boldsymbol{\chi} = \mathbf{x_1} - \mathbf{x_2}$.

We will make two simplifying assumptions in the analysis of the system performance. First, we assume that the interleaver π is a uniform interleaver, as defined by Benedetto *et al.* [13].

Definition 1: A uniform interleaver of length k is a probabilistic device which maps a given input word of weight w into all $\binom{k}{w}$ distinct permutations of it with equal probability $1/\binom{k}{w}$.

The uniform interleaver may be thought of as the average over the ensemble of all deterministic length-k interleavers (i.e., permutations), assuming a uniform distribution. The use of this device has proven to be very valuable in analyzing the average ML performance of parallel and serial concatenated coding architectures.

Second, we make the assumption that, for any error word $\mathbf{e}$, the contribution to $\overline{T}(d_E)$ of all error words $\boldsymbol{\epsilon} = \mathbf{b_1} - \mathbf{b_2}$, where $\mathbf{b_1} = \mathbf{b_2} \oplus \mathbf{e}$, is approximately equal to the contribution of the set of error words produced when $\mathbf{b_1}$ and $\mathbf{b_2}$ are not restricted to lie within the code. This is equivalent to treating the permuted code bits within an error event at the output of the interleaver as independent and identically distributed (i.i.d.), with equiprobable bit values. The resulting estimate of the contribution to the Euclidean weight enumerator therefore depends only upon the Hamming weight of the modulo-2 error word $\mathbf{e}$, rather than the specific signed error words $\boldsymbol{\epsilon}$ produced by the actual codeword differences. The rationale behind this second assumption is that the system uses a very high rate, linear encoder, in tandem with the uniform interleaver.

In Section II-A, we investigate the relationship between the Hamming weight $d_H(\mathbf{f})$ of the interleaved Hamming error word and the squared Euclidean weight $d_E^2(\boldsymbol{\zeta}) = \sum_{i=1}^{N} \chi_i^2$ of the corresponding output error word on the dicode channel, $h(D) = 1 - D$. We then examine the distribution of the number of error events induced by the action of the uniform interleaver. Using these results, we then derive an estimate for $\overline{T}(d_E)$ and the system performance.

In Section II-B, we derive the corresponding result for the dicode channel with the precoder $g(D) = 1/(1 \oplus D)$.

In Section II-C, we extend the results in Sections II-A and II-B to the PR4 channel.

A. Dicode Channel with No Precoder

1) Error Event Distance Properties: Fig. 2 shows a trellis section for the dicode channel with no precoder. The branch labels are of the form c_i/x_i, where c_i is the input to the channel at time i, and x_i is the corresponding channel output.

Let $\mathbf{f}$ be an error word with Hamming weight $l = d_H(\mathbf{f})$, corresponding to a possibly compound input error event. Referring

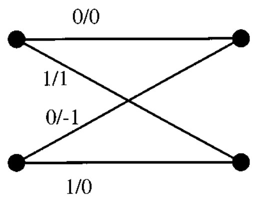

Fig. 2. Trellis section for the dicode channel.

to Fig. 2, and assuming a fixed initial state for all codewords, $\mathbf{f}$ can be uniquely decomposed into a concatenation of disjoint error subevents $\mathbf{f_i}$, $i = 1, \ldots, m$, for some $m \geq 1$, consisting of one or more consecutive errors. Letting $l_i = d_H(\mathbf{f_i})$ denote the number of errors in the subevent, we have $l = \sum_{i=1}^{m} l_i$. For $i \leq m-1$, the subevent $\mathbf{f_i}$ corresponds to a simple closed error event on the trellis, diverging from and remerging with the correct path, with no common intermediate states. For $i = m$, the corresponding subevent $\mathbf{f_m}$ may be either closed or open; in the latter case, the paths diverge and never remerge.

In Fig. 2, it can be seen that diverging and remerging branches correspond to a squared Euclidean distance of 1. Parallel branches have Euclidean distance 0 and crossing branches have squared Euclidean distance 4. We further note that crossing branches at time k within an error event occur when the input symbol c_k differs from the previous input symbol c_{k-1}.

Let j_i denote the bit position at which the subevent $\mathbf{f_i}$ begins. If $c_k \oplus c_{k-1} = 1$ for $j_i < k < j_i + l_i$, then the branches at time k are crossing with squared Euclidean distance 4. For a closed subevent, the contribution to the squared Euclidean distance is given by

$$d_E^2(\mathbf{f_i}) = 2 + 4 \sum_{k=j_i+1}^{j_i+l_i-1} c_k \oplus c_{k-1} \tag{5}$$

where the first term is the contribution from the diverging and remerging branches, and the second term is the contribution from crossing branches within the subevent. If $\mathbf{f_m}$ is open, there is no remerging branch, and the contribution is

$$d_E^2(\mathbf{f_m}) = 1 + 4 \sum_{k=j_m+1}^{j_m+l_m-1} c_k \oplus c_{k-1}. \tag{6}$$

The compound error event $\mathbf{f}$ generates squared Euclidean distance

$$\begin{aligned} d_E^2(\mathbf{f}) &= \sum_{i=1}^{m} d_E^2(\mathbf{f_i}) \\ &= 2m + 4 \sum_{i=1}^{m} \sum_{k=j_i+1}^{j_i+l_i-1} c_k \oplus c_{k-1} - \delta(j_m + l_m - 1 - N) \end{aligned} \tag{7}$$

where

$$\delta(n) = \begin{cases} 1, & \text{if } n = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

The expression in (7) requires knowledge of the distribution of the input symbols $\mathbf{c}$. Invoking the assumption regarding the distribution of code bit values in the error events—namely, that their values are i.i.d. and equiprobable—we obtain an approximate contribution of an error word $\mathbf{f}$ of Hamming weight d to the average Euclidean distance spectrum as follows. Denote the number of crossing branches within subevents as y. For a given set of d and m, there are a total of 2^{d-m} different paths within subevents. There are $\binom{d-m}{y}$ different paths with y crossing branches, and the distribution of the number of crossing branches is

$$\Pr(y|d, m) = \binom{d-m}{y} 0.5^{d-m}. \tag{9}$$

Therefore, for the case when $\mathbf{f_m}$ is closed, we obtain

$$\begin{aligned} &\Pr(d_E | d, m, \mathbf{f_m} \text{ closed, i.i.d.}) \\ &= \begin{cases} \binom{d-m}{(d_E^2 - 2m)/4} 0.5^{d-m}, & \text{if } (d_E^2 - 2m)/4 \text{ is an integer} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{10}$$

as the probability that an erroneous codeword is at Euclidean distance d_E from the correct codeword, conditioned on Hamming weight d, m error subevents where the last event is closed, and independent equiprobable bit values within error subevents. Similarly, when $\mathbf{f_m}$ is open, we have

$$\begin{aligned} &\Pr(d_E | d, m, \mathbf{f_m} \text{ open, i.i.d.}) \\ &= \begin{cases} \binom{d-m}{(d_E^2 - 2m+1)/4} 0.5^{d-m}, & \text{if } (d_E^2 - 2m+1)/4 \text{ is an integer} \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{11}$$

These distributions can be used to evaluate the Euclidean distance between channel output sequences corresponding to codewords at Hamming distance d, when the error words consist of m subevents.

2) Subevent Distribution at the Interleaver Output: Let $\mathbf{e}$ be an error word with Hamming weight $d_H(\mathbf{e}) = d$. A specified interleaver will map $\mathbf{e}$ into an error word $\mathbf{f}$ which can be decomposed into m error subevents, $\mathbf{f_i}$, $i = 1, \ldots, m$, with corresponding weights l_i satisfying $l = \sum_{i=1}^{m} l_i$, as described above. A uniform interleaver maps the error word $\mathbf{e}$ into all $\binom{N}{d}$ possible error words $\mathbf{f}$ with equal probability $1/\binom{N}{d}$. In this section, we determine the distribution of the number m of subevents of $\mathbf{f}$, conditioned upon the error word $\mathbf{e}$ having Hamming weight d, under the action of the uniform interleaver.

There are $\binom{d-1}{m-1}$ distinct decompositions of a sequence of d elements into m subsequences, each of length at least 1. The number of configurations in which these m subsequences can occur in a word of length N, with consecutive subsequences separated by at least one position, is given by $\binom{N-d+1}{m}$, so there are $\binom{N-d+1}{m}\binom{d-1}{m-1}$ weight d words with m subevents. Taking into consideration the nature of the subevent $\mathbf{f_m}$, we can compute the conditional joint probabilities

$$\Pr(m,\, \mathbf{f_m} \text{ closed}|d) = \frac{\binom{N-d}{m}\binom{d-1}{m-1}}{\binom{N}{d}} \tag{12}$$

and

$$\Pr(m,\, \mathbf{f_m} \text{ open}|d) = \frac{\binom{N-d}{m-1}\binom{d-1}{m-1}}{\binom{N}{d}}. \tag{13}$$

3) Approximation of the Euclidean Weight Enumerator: If we define $\Pr(d_E|d)$ as the conditional probability that the Euclidean distance between two codewords is d_E, given that the Hamming distance is d, then the average Euclidean weight enumerator $\overline{T}(d_E)$ is given by

$$\begin{aligned}\overline{T}(d_E) &= \sum_{d=1}^{N} A(d)\, \Pr(d_E|d) \\ &= \sum_{d=1}^{N} A(d) \sum_{m=1}^{d} \\ &\quad \cdot (\Pr(d_E|d,\, m,\, f_m \text{ closed}) \Pr(m,\, f_m \text{ closed}|d) \\ &\quad + \Pr(d_E|d,\, m,\, f_m \text{ open}) \Pr(m,\, f_m \text{ open}|d)).\end{aligned} \tag{14}$$

The approximation, denoted $\tau(d_E)$, is then given by substituting the approximations given in (10) and (11), along with the conditional joint probabilities given in (12) and (13), into (14), yielding

$$\begin{aligned}\tau(d_E) = \sum_{d=1}^{N} A(d) \frac{1}{\binom{N}{d}} &\cdot \Bigg(\sum_{m: d_E^2 - 2m = 0 \bmod (4)} \binom{d-m}{(d_E^2-2m)/4} \\ &\cdot 0.5^{d-m} \binom{N-d}{m} \binom{d-1}{m-1} \\ &+ \sum_{m: d_E^2 - 2m + 1 = 0 \bmod (4)} \binom{d-m}{(d_E^2-2m+1)/4} \\ &\cdot 0.5^{d-m} \binom{N-d}{m-1} \binom{d-1}{m-1} \Bigg).\end{aligned} \tag{15}$$

The approximate average information Hamming distance to codewords at Euclidean distance d_E, denoted $\omega(d_E)$, is similarly computed by substitution into

$$\overline{w}(d_E) = \frac{1}{\overline{T}(d_E)} \sum_{d=1}^{N} A(d) \overline{W}(d) \Pr(d_E|d) \tag{16}$$

where $\overline{W}(k)$ is the average input weight for codewords of Hamming weight $d_H = k$.

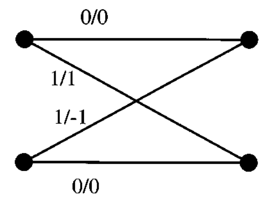

Fig. 3. Trellis section for the precoded dicode channel.

B. Precoded Dicode

1) Error Event Distance Properties: Fig. 3 shows a trellis section for the dicode channel with precoder $g(D) = 1/(1 \oplus D)$. The branch labels are of the form c_i/x_i, where c_i is the input to the precoder at time i, and x_i is the corresponding channel output.

Referring to Fig. 3, it can be seen that an error word $\mathbf{f}$ may be decomposed into a sequence of $m = \lceil d_H(\mathbf{f})/2 \rceil$ simple error subevents $\mathbf{f_i}$, $i = 1, \ldots, m$. For $1 \le i \le m-1$, each subevent is closed. Subevent $\mathbf{f_m}$ may be either closed or open. The length of the subevent $\mathbf{f_i}$ is denoted l_i, and the Hamming weight of a subevent satisfies

$$d_H(\mathbf{f_i}) = \begin{cases} 2, & \text{if } i = 1, \ldots, m-1 \\ 2, & \text{if } i = m \text{ and } d_H(\mathbf{f}) \text{ even} \\ 1, & \text{if } i = m \text{ and } d_H(\mathbf{f}) \text{ odd.} \end{cases} \tag{17}$$

Let j_i^0 denote the bit position in the word where error event $\mathbf{f_i}$ begins. For closed events, let j_i^1 denote the bit position where it terminates. Then, $l_i = j_i^1 - j_i^0 + 1$ for all closed subevents. If $\mathbf{f_m}$ is open, we define $j_m^1 = N+1$, and $l_m = j_m^1 - j_m^0$. Finally, we define $L = \sum_{i=1}^{m} l_i$ as the total error event length.

As for the nonprecoded case, diverging and remerging branches contribute a squared Euclidean distance of 1. Parallel branches contribute distance 0 and crossing branches contribute squared Euclidean distance 4. Crossing branches correspond to the input symbol at that time being 1. Therefore, the total contribution $d_E^2(\mathbf{f_i})$ of a subevent $\mathbf{f_i}$ to the squared Euclidean distance at the channel output is given by

$$d_E^2(\mathbf{f_i}) = d_H(\mathbf{f_i}) + 4 \sum_{k=j_i^0+1}^{j_i^1-1} c_k. \tag{18}$$

The error word $\mathbf{f}$ has total squared Euclidean distance

$$d_E^2(\mathbf{f}) = \sum_{i=1}^{m} d_E^2(\mathbf{f_i}) = d_H(\mathbf{f}) + 4 \sum_{i=1}^{m} \sum_{k=j_i^0+1}^{j_i^1-1} c_k. \tag{19}$$

Thus, the squared Euclidean distance between two codewords is equal to the Hamming distance plus four times the number of ones within error subevents.

Invoking, as in the nonprecoded case, the assumption regarding the distribution of code bit values in the error events—namely, that their values are i.i.d. and equiprobable—we obtain an approximate contribution of an error word $\mathbf{f}$ of Hamming weight d to the average Euclidean weight enumerator. Under this assumption, we note that the distribution of the squared Euclidean distance is a function of the distribution of the number of ones within error subevents. Denote by y_1 the number of ones within error subevents. For error events of total length L and Hamming weight d, the distribution of y_1 is given by

$$\Pr(y_1|d, L) = \binom{L-d}{y_1} 0.5^{L-d}. \tag{20}$$

For the probability that an erroneous codeword is at Euclidean distance d_E from the correct codeword, conditioned on the Hamming distance being d, the total length of the suberror events, we obtain

$$\Pr(d_E|d, L, \text{ i.i.d.}) = \begin{cases} \binom{L-d}{(d_E^2-d)/4} 0.5^{L-d}, & \text{if } (d_E^2-d)/4 \text{ is an integer} \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

The i.i.d. approximation is justified for error words $\mathbf{f}$ with a small value of L by the action of the uniform interleaver. For error words with large value of L, the contribution to the dominant terms of the Euclidean weight enumerator will be negligible, in any case, due to the low probability of then generating small Euclidean distance.

2) Subevent Distribution at the Interleaver Output: Let $\mathbf{e}$ be an error word of Hamming weight d. A permuted error word $\mathbf{f}$ can be decomposed into $m = \lceil d/2 \rceil$ error events $\mathbf{f_i}$, as described in Section II-B-1. In this section, we determine the conditional distribution of the total length L of subevents generated by the action of a uniform interleaver upon error words $\mathbf{e}$ of Hamming weight d.

The distribution is computed in two steps. First, we find the number of unique back-to-back concatenations of m subevents of total length L. Then, we determine the number of configurations in which the m subevents can occur in a word of length N.

Consider the following description of the permuted error word $\mathbf{f}$

$$0, \ldots, 0, 1_1, 0_1, \ldots, 0_1, 1_1, 0, \ldots, 0, \ldots, 1_m, 0_m, \ldots, 0_m, 1_m, 0, \ldots, 0$$

where the subscript denotes to which subevent a bit belongs. There are $\binom{L-\lceil (d-1)/2 \rceil - 1}{\lfloor (d-1)/2 \rfloor}$ unique back-to-back concatenations of subevents $\mathbf{f_i}$ of total length L. If d is even, the remaining $N-L$ bits can be partitioned in $\binom{N-L+m}{m}$ different ways. If d is odd, the permutation has to end with an open error event, so there are $\binom{N-L+m-1}{m-1}$ possible permutations.

The conditional distribution $\Pr(L|d)$ of the total length L, given an error word of Hamming weight $d_H(e) = d$, is therefore given by

$$\Pr(L|d) = \frac{\binom{N-L+\lfloor d/2 \rfloor}{\lfloor d/2 \rfloor} \binom{L-1-\lceil (d-1)/2 \rceil}{\lfloor (d-1)/2 \rfloor}}{\binom{N}{d}}. \tag{22}$$

Remark: Divsalar *et al.* derived a similar expression for the input–output weight enumerator for the accumulate code in the context of repeat-accumulate codes [14, eq. (5.3)].

3) Approximation of the Euclidean Weight Enumerator: The approximate Euclidean weight enumerator $\tau(d_E)$ can be computed by substituting (21) and (22) into

$$\overline{T}(d_E) = \sum_{d=1}^{N} A(d) \sum_{L=d}^{N} \Pr(d_E|d, L) \Pr(L|d). \tag{23}$$

In a similar way, the approximate average input error weight enumerator may be obtained by appropriate substitutions into

$$\overline{w}(d_E) = \frac{1}{\overline{T}(d_E)} \sum_{d=1}^{N} A(d)\overline{W}(d) \sum_{L=d}^{N} \Pr(d_E|d, L) \Pr(L|d). \tag{24}$$

C. Extension to PR4

By observing that the PR4 channel can be viewed as two interleaved dicode channels, the extension of the results to PR4 is straightforward. The details are given in Appendix A.

We note that Duman and Kurtas [15] have utilized the i.i.d. assumption to derive performance estimates for higher order partial response channels. However, as a result of the increased complexity involved in computing the error events for the partial response channel, they have resorted to using the transfer function matrix approach; see, for example, [12] and [16].

III. Computed Bounds and Simulation Results

In this section, we compute truncated ML union bound estimates for the turbo-equalized dicode channel using the method described above, and we compare these with computer simulation results obtained using iterative decoding. Although suboptimal, the iterative decoder should be comparable in performance to the ML decoder once the SNR reaches a moderately high value.

We consider two outer encoders as follows: 1) a rate 1/2, recursive systematic convolutional (RSC) encoder with encoder polynomials $(31, 33)_{\text{octal}}$, with parity bits punctured to yield code rate 8/9, and 2) a turbo code consisting of a parallel concatenation of two of the RSC encoders with parity bits punctured to achieve rate 4/5. Both encoders use an information block of size $K = 4096$. The iterative decoder used in the simulations incorporates *a posteriori* probability (APP) decoders for both the channel and the component codes. Soft information is shared between all decoders for up to ten full iterations. If three consecutive iterations generate the same sequence estimate, then the iterations are terminated in order to reduce the simulation time. Fig. 4 shows the WER results for the rate 8/9 system, with and without precoder. Fig. 5 shows the corresponding results for the BER.

In Table I, we show the Euclidean weight enumerator estimates, obtained from (15) and (23), for the system using an outer convolutional code with no channel precoder, as well as with a channel precoder. We note that the dominant contributor to the estimated error rate is determined by the Euclidean distance and corresponding multiplicity that together yield the largest spectral component in the union bound. For the dicode channel, this will, at moderate SNR, be the $d_E^2 = 4$ component, and it is the result of Hamming weight-2 error words that are mapped

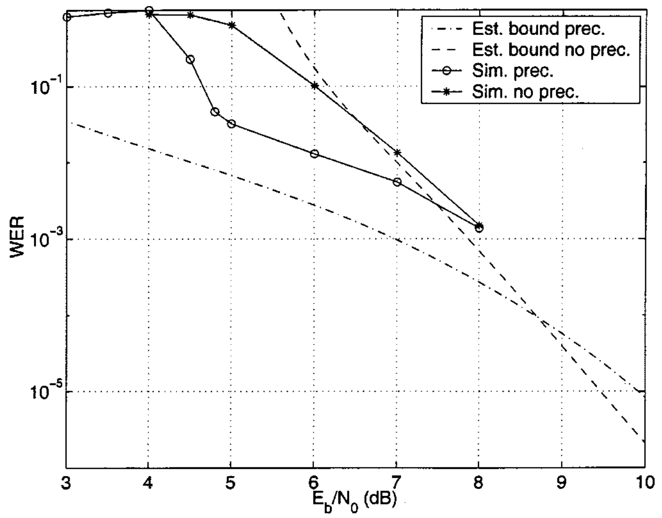

Fig. 4. WER union bound estimates and simulation results for outer convolutional code.

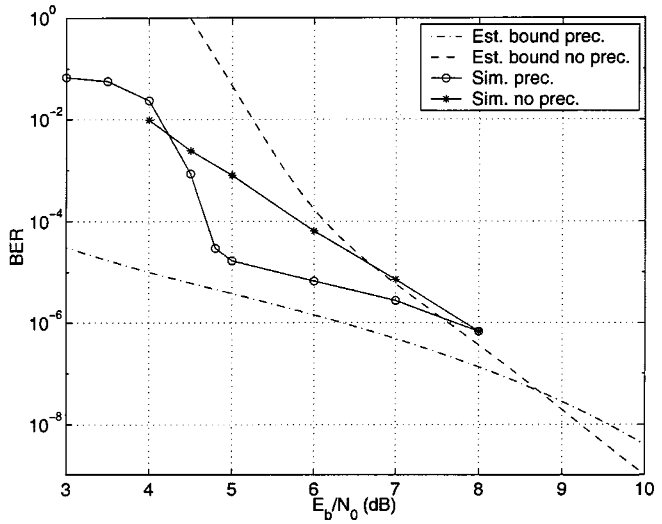

Fig. 5. BER union bound estimates and simulation results for outer convolutional code.

TABLE I
HAMMING AND APPROXIMATE EUCLIDEAN WEIGHT ENUMERATORS FOR SYSTEMS WITH OUTER CONVOLUTIONAL CODE

Outer Code			Precoded	Not Precoded
d_H	$\overline{A}(d_H)$	d_E^2	$\tau(d_E)$	$\tau(d_E)$
2	510	2	0.4426	0.1122
3	21421	3	0.02421	0.2274
4	357864	4	0.8084	523.6
5	13192299	5	0.06468	14.25
6	389079383	6	4.255	21864
7	9010184299	7	0.3166	336.1
8	236369355044	8	15.68	386587

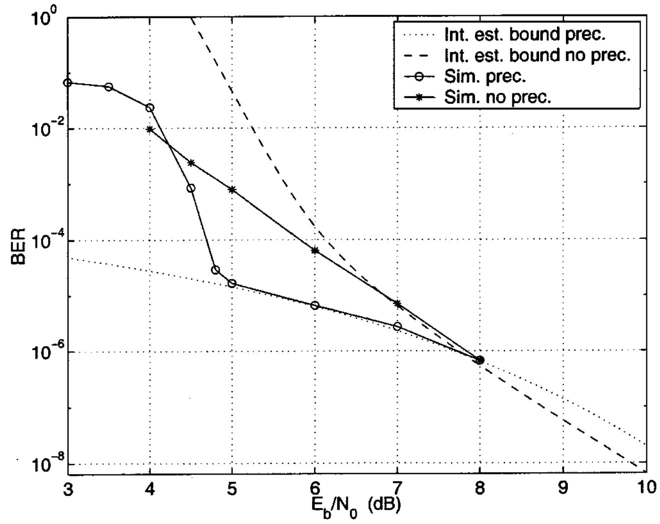

Fig. 6. Truncated bound estimates for interleaver used versus simulation result for outer convolutional code.

into two subevents, each giving squared Euclidean distance 2. For the precoded dicode channel, it is the $d_E^2 = 2$ component, and it is also the result of Hamming weight-2 error words that are mapped to two adjacent positions, or under certain circumstances, to positions separated by a few bits. Since odd values of the squared Euclidean distance are the result of open events, they are not as common as even values.

The randomly selected interleavers used in the simulations and the uniform interleaver induce different weight enumerators. Therefore, the estimated bounds and the simulation results differ, and in the case without the precoder, the simulation curve crosses the bound curve. We investigated Hamming weight-2 error events for the specific interleaver used in the simulations. For the nonprecoded case, we adjusted the parts of the estimated distance spectrum corresponding to Hamming weight-2 error events to reflect the mappings by the actual interleaver used. The punctured code has, as shown in Table I, 510 codewords at Hamming distance 2 from each codeword. The interleaver that was used maps two of these events in such a manner that the erroneous bits are adjacent. Each of those error events will give one error subevent, and the squared Euclidean distance is 2 with probability 1/2 and 6 with probability 1/2. This means that Hamming weight-2 error events will contribute 1 to the approximate average Euclidean weight enumerators $\tau(\sqrt{2})$ and $\tau(\sqrt{6})$. Therefore, we will have $\tau(\sqrt{2}) \geq 1$, a value which is nine times larger than the corresponding term for the uniform interleaver. Each of the remaining 508 Hamming weight-2 error events are mapped into two subevents. All of these events generate squared Euclidean distance $d_E^2 = 4$ events, with the exception of one such event whose last subevent is open and which, therefore, generates only squared Euclidean distance $d_E^2 = 3$. It follows that $\tau(\sqrt{3}) \geq 1$.

For the precoded case, we determined the distance spectrum for Hamming weight-2 error events of length less than 16 at the output of the interleaver. We note that for these events $\tau(\sqrt{2}) = 3.31$, which is about eight times greater than the corresponding value for the uniform interleaver.

Fig. 6 compares the BER simulation results with the BER estimates for the interleaver used in the simulations by applying the union bound to the distance spectrum terms obtained above. The fit between the analysis and simulation is improved, particularly in the precoded case. The difference in estimated performance for the uniform interleaver and the interleaver used in the simulations indicates that the choice of interleaver can influence performance significantly.

TABLE II
BER WITH PRECODER AT $E_b/N_0 = 6.0$ dB FOR OUTER CONVOLUTIONAL CODES AND FOUR DIFFERENT INTERLEAVERS

Interleaver	1	2	3	4
BER (6.0 dB) $/10^{-5}$	0.6610	0.2689	0.0766	0.0407

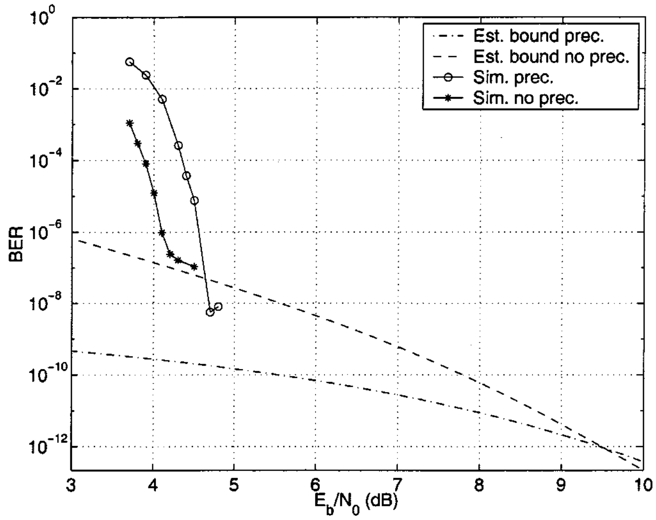

Fig. 7. BER bound estimates versus simulation results for precoded and nonprecoded turbo system.

TABLE III
APPROXIMATE HAMMING AND EUCLIDEAN WEIGHT ENUMERATORS FOR TURBO-CODED SYSTEMS

Outer Code d_H	$\overline{A}(d_H)$	d_E^2	Precoded $\tau(d_E)$	Not Precoded $\tau(d_E)$
2	0.031	2	0.000024	0.000006
3	0.462	3	0.000000	0.000000
4	2.111	4	0.000004	0.031273
5	4.100	5	0.000000	0.000000
6	8.842	6	0.000024	0.464084
7	20.337	7	0.000001	0.000000
8	50.743	8	0.000008	2.114827

The impact of the interleaver in the precoded case is further reflected in simulation results for three additional, randomly selected interleavers. Table II shows the BER values at $E_b/N_0 = 6$ dB for the interleaver used to generate the results in Fig. 6, followed by the additional three interleavers. The table suggests that suitable interleaver design can significantly improve the system performance.

Fig. 7 shows analytical BER estimates and simulation results for the rate 4/5 turbo-coded systems. The Hamming weight enumerator and the estimated Euclidean weight enumerator for the turbo-coded system are shown in Table III. The bound for the precoded system is much lower at SNR up to about $E_b/N_0 = 9.5$ dB. However, in simulations, the system without precoder is superior down to $P_b \approx 2 \cdot 10^{-7}$, at which point the simulated BER curve flattens out and tends to follow the analytical curve. In fact, above 4.7 dB, the precoded system becomes superior to the system without the precoder, as predicted by the analysis. The explanation for the behavior observed at very low SNR remains an open issue.

IV. CONCLUSIONS

We have presented an analytical method for estimating the average Euclidean distance spectrum for a serially concatenated, trellis-coded partial response channel. The technique was applied to the dicode channel, with and without precoding, and was extended to the PR4 channel. Using truncated union bounds, we derived analytical BER and WER results and compared them to computer simulations. The analytical results indicate that the precoded systems should perform better in the floor region, as was empirically confirmed. Future research directions are to bound the effect of the i.i.d. assumption, develop exact methods for higher-order channels, and include the entire Hamming weight spectrum of the outer code in the computations.

APPENDIX
EXTENSION TO PR4

A. PR4 Channel with No Precoder

For the nonprecoded case we observe that the conditional distributions for the Euclidean weight contribution from an error word $\mathbf{f}$ of Hamming weight d as given in (10) and (11) can be modified as follows:

$$\Pr(d_E|d,\, m,\, \mu \text{ open events, i.i.d.}) = \begin{cases} \binom{d-m}{(d_E^2-2m+\mu)/4} 0.5^{d-m}, & \text{if } \binom{d-m}{(d_E^2-2m+\mu)/4} \text{ is an integer} \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

where $\mu \in \{0, 1, 2\}$ denotes the number of open events. We note that for the dicode channel, μ can only have the values 0 and 1, but that for the PR4 channels there can be an open event for both interleaves and therefore μ can also have the value 2.

For the conditional joint probabilities $\Pr(m, \mu|d)$, we have to take into consideration the possible decompositions of m and d into the two interleaves. Assume $m = m_1 + m_2$, where m_1 and m_2 denote the number of subevents in the first and second interleave, respectively. In a similar way, we split up the Hamming weight into $d = d_1 + d_2$ and $\mu = \mu_1 + \mu_2$. By taking the sum over all possible partitions of m into m_1 and m_2, and d into d_1 and d_2, we have for each μ

$$\begin{aligned} \Pr(m,\, \mu=0|d) &= \sum_{\substack{d_1=1 \\ d_2=d-d_1}}^{d-1} \Pr(d_1|d) \sum_{\substack{m_1=1 \\ m_2=m-m_1}}^{\min(m-1,\, d_1)} \\ &\quad \cdot \Pr(m_1,\, \mu_1=0|d_1) \Pr(m_2,\, \mu_2=0|d_2) \\ &\quad + \Pr(d_1=d|d) \Pr(m,\, \mu_1=0|d) \\ &\quad + \Pr(d_2=d|d) \Pr(m,\, \mu_2=0|d) \\ &= \sum_{d_1=1}^{d-1} \Pr(d_1|d) \sum_{m_1=1}^{\min(m-1,\, d_1)} \Pr(m_1,\, \mu_1=0|d_1) \\ &\quad \cdot \Pr(m-m_1,\, \mu_2=0|d-d_1) \\ &\quad + \Pr(d_1=d|d) \Pr(m,\, \mu_1=0|d) \\ &\quad + \Pr(d_2=d|d) \Pr(m,\, \mu_2=0|d) \end{aligned} \quad (26)$$

$$\begin{aligned}\Pr(m,\,\mu=1|d) &= 2\sum_{d_1=1}^{d-1}\Pr(d_1|d)\sum_{m_1=1}^{\min(m-1,\,d_1)}\Pr(m_1,\,\mu_1=1|d_1)\\ &\quad\cdot\Pr(m-m_1,\,\mu_2=0|d-d_1)\\ &\quad+2\,\Pr(d_1=d|d)\Pr(m,\,\mu_1=1|d) \qquad (27)\end{aligned}$$

$$\begin{aligned}\Pr(m,\,\mu=2|d) &= \sum_{d_1=1}^{d-1}\Pr(d_1|d)\sum_{m_1=1}^{\min(m-1,\,d_1)}\Pr(m_1,\,\mu_1=1|d_1)\\ &\quad\cdot\Pr(m-m_1,\,\mu_2=1|d-d_1). \qquad (28)\end{aligned}$$

We define the length n of each interleave as $n = N/2$. The probability that the first interleave has Hamming weight d_1 conditioned on total Hamming weight d is given by

$$\Pr(d_1|d) = \frac{\binom{n}{d_1}\binom{n}{d-d_1}}{\binom{N}{d}}. \qquad (29)$$

We also note that the expressions for the joint conditional probabilities for the number of subevents in each interleave are similar to (12) and (13), namely

$$\Pr(m_i,\,\mu_i|d_i) = \begin{cases}\dfrac{\binom{n-d_i}{m_i}\binom{d_i-1}{m_i-1}}{\binom{n}{d_i}}, & \mu_i=0\\[2ex] \dfrac{\binom{n-d_i}{m_i-1}\binom{d_i-1}{m_i-1}}{\binom{n}{d_i}}, & \mu_i=1.\end{cases} \qquad (30)$$

By inserting (29) and (30) into (26)–(28), we have

$$\begin{aligned}\Pr(m,\,\mu=0|d) &= \sum_{d_1=1}^{d-1}\frac{1}{\binom{N}{d}}\sum_{m_1=d_1}^{m-d+d_1}\binom{n-d_1}{m_1}\binom{d_1-1}{m_1-1}\binom{n-d+d_1}{m-m_1}\\ &\quad\cdot\binom{d-d_1-1}{m-m_1-1}+2\,\frac{\binom{n-d}{m}\binom{d-1}{m-1}}{\binom{N}{d}} \qquad (31)\end{aligned}$$

$$\begin{aligned}\Pr(m,\,\mu=1|d) &= \sum_{d_1=1}^{d-1}\frac{2}{\binom{N}{d}}\sum_{m_1=d_1}^{m-d+d_1}\binom{n-d_1}{m_1-1}\binom{d_1-1}{m_1-1}\binom{n-d+d_1}{m-m_1}\\ &\quad\cdot\binom{d-d_1-1}{m-m_1-1}+2\,\frac{\binom{n-d}{m-1}\binom{d-1}{m-1}}{\binom{N}{d}} \qquad (32)\end{aligned}$$

and

$$\begin{aligned}\Pr(m,\,\mu=2|d) &= \sum_{d_1=1}^{d-1}\frac{1}{\binom{N}{d}}\sum_{m_1=d_1}^{m-d+d_1}\binom{n-d_1}{m_1-1}\binom{d_1-1}{m_1-1}\binom{n-d+d_1}{m-m_1-1}\\ &\quad\cdot\binom{d-d_1-1}{m-m_1-1}. \qquad (33)\end{aligned}$$

We modify (14) as follows:

$$\begin{aligned}\overline{T}(d_E) &= \sum_{d=1}^{N}A(d)\Pr(d_E|d)\\ &= \sum_{d=1}^{N}A(d)\sum_{m=1}^{d}\\ &\quad\cdot[\Pr(d_E|d,\,m,\,\mu=0)\Pr(m,\,\mu=0|d)\\ &\quad+\Pr(d_E|d,\,m,\,\mu=1)\Pr(m,\,\mu=1|d)\\ &\quad+\Pr(d_E|d,\,m,\,\mu=2)\Pr(m,\,\mu=2|d)]. \qquad (34)\end{aligned}$$

The approximate Euclidean weight enumerator, denoted $\tau(d_E)$, is then given by substituting the approximations given in (25), along with the conditional joint probabilities given in (31)–(33), into (34).

B. PR4 Channel with Precoder

When the precoder $1/(1\oplus D^2)$ is used for the PR4 channel, we easily obtain an expression for the conditional probability $\Pr(L|d)$ corresponding to (22). We again view the precoded PR4 channel as two interleaved, precoded dicode channels, each with the same conditional distribution as (22), but with the length of the interleaver being $n = N/2$ instead. Thus, we have

$$\Pr(L_i|d_i) = \frac{\binom{n-L_i+\lfloor d_i/2\rfloor}{\lfloor d_i/2\rfloor}\binom{L_i-1-\lfloor d_i/2\rfloor}{\lfloor (d_i-1)/2\rfloor}}{\binom{n}{d_i}} \qquad (35)$$

where L_i is the total length of the subevents in interleave i, and d_i is the Hamming weight of interleave i. Note that $L = L_1+L_2$ and $d = d_1 + d_2$. We write

$$\begin{aligned}\Pr(L|d) &= \frac{1}{\binom{N}{d}}\left[\sum_{d_1=1}^{d-1}\sum_{L_1=d_1}^{L-d+d_1}\binom{n-L_1+\lfloor d_1/2\rfloor}{\lfloor d_1/2\rfloor}\right.\\ &\quad\cdot\binom{L_1-1-\lfloor d_1/2\rfloor}{\lfloor (d_1-1)/2\rfloor}\\ &\quad\cdot\binom{n-L+L_1+\lfloor (d-d_1)/2\rfloor}{\lfloor (d-d_1)/2\rfloor}\\ &\quad\left.\cdot\binom{L-L_1-1-\lfloor (d-d_1)/2\rfloor}{\lfloor (d-d_1-1)/2\rfloor}\right]\\ &\quad+\frac{2}{\binom{N}{d}}\binom{n-L+\lfloor d/2\rfloor}{\lfloor d/2\rfloor}\binom{L-1-\lfloor d/2\rfloor}{\lfloor (d-1)/2\rfloor}. \qquad (36)\end{aligned}$$

The approximate Euclidean weight enumerator $\tau(d_E)$ can be computed by inserting (21) and (36) into (23).

ACKNOWLEDGMENT

The authors would like to thank T. Souvignier and H. Pfister for helpful insights and interesting discussions, and the reviewers and Associate Editor for their detailed comments and suggestions. The authors would also like to thank D. N. Rowitch [16] for providing the necessary Hamming weight enumerators.

REFERENCES

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Communications*, vol. 2, Geneva, Switzerland, May 1993, pp. 1064–1070.

[2] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[3] S. Benedetto, G. Montorsi, D. Divsalar, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *Telecom. Data Acquisition Prog. Rep.*, vol. 42, pp. 1–26, Aug. 1996.

[4] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507–511, Sept.–Oct. 1995.

[5] C. Heegard, "Turbo coding for magnetic recording," in *Proc. IEEE Information Theory Workshop*, San Diego, CA, USA, Feb. 1998, pp. 18–19.

[6] W. Pusch, D. Weinrichter, and M. Taferner, "Turbo-codes matched to the $1 - D^2$ partial response channel," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, Aug. 1998, p. 62.

[7] M. C. Reed and C. B. Schlegel, "An iterative receiver for the partial response channel," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, Aug. 1998, p. 63.

[8] W. E. Ryan, L. L. McPheters, and S. W. McLaughlin, "Combined turbo coding and turbo equalization for PR4-equalized Lorentzian channels," in *Proc. Conf. Information Sciences and Systems*, Mar. 1998.

[9] W. E. Ryan, "Performance of high rate turbo codes on a PR4-equalized magnetic recording channel," in *Proc. IEEE Int. Conf. Communications*, Atlanta, GA, June 1998, pp. 947–951.

[10] T. Souvignier, A. Friedman, M. Öberg, P. H. Siegel, R. E. Swanson, and J. K. Wolf, "Turbo codes for PR4: Parallel versus serial concatenation," in *Proc. IEEE Int. Conf. Communications*, Vancouver, BC, Canada, June 1999, pp. 1638–1642.

[11] L. L. McPheters, S. V. McLaughlin, and K. R. Narayanan, "Precoded PRML, serial concatenation and iterative (turbo) decoding for digital magnetic recording," in *Proc. IEEE Int. Magnetics Conf.*, Kyongju, Korea, May 1999.

[12] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[13] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 409–428, Mar. 1996.

[14] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for "turbo-like" codes," in *Proc. 35th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sept. 1998, pp. 201–210.

[15] T. M. Duman and E. Kurtas, "Performance bounds for high rate linear codes over partial response channels," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 258.

[16] D. N. Rowitch, "Convolutional and turbo coded multicarrier direct sequence CDMA, and applications of turbo codes to hybrid ARQ communication systems," PhD dissertation, Univ. California, San Diego, La Jolla, June 1998.

Mats Öberg (S'97–M'01) was born in Piteå, Sweden, in 1969. He received the M.S. degree in electrical engineering from Lund University, Lund, Sweden, in 1995. He received the Ph.D. degree in electrical engineering (communication theory and systems) from the University of California, San Diego, La Jolla, in 2000.

He is a Staff Design Engineer at Marvell Semiconductor, Inc., Sunnyvale, CA. His research interests include coding and signal processing for wireless communications and magnetic storage applications.

Paul H. Siegel (M'82–SM'90–F'97) received the S.B. degree in mathematics in 1975 and the Ph.D. degree in mathematics in 1979, both from the Massachusetts Institute of Technology, Cambridge. He held a Chaim Weizmann Fellowship during a year of postdoctoral study at the Courant Institute, New York University.

He was with the IBM Research Division from 1980 to 1995. He joined the Faculty of the School of Engineering at the University of California, San Diego in July 1995, where he is currently Professor of Electrical and Computer Engineering. He is affiliated with the Center for Wireless Communications and became Director of the Center for Magnetic Recording Research in September 2000. He holds several patents in the area of coding and detection for digital recording systems. His primary research interest is the mathematical foundations of signal processing and coding, especially as applicable to digital data storage and communications.

Prof. Siegel is a member of Phi Beta Kappa. He was a member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996. He served as Co-Guest Editor of the May 1991 Special Issue on Coding for Storage Devices of the IEEE TRANSACTIONS ON INFORMATION THEORY, and served as Associate Editor for Coding Techniques from 1992 to 1995. He was co-recipient of the 1992 IEEE Information Theory Society Paper Award and the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award.