

The Error-Pattern-Correcting Turbo Equalizer: Spectrum Thinning at High SNRs

Hakim Alhussien, *Member, IEEE*, and Jaekyun Moon, *Fellow, IEEE*

Dedicated to the memory of Ralf Koetter (1963–2009)

Abstract—The error-pattern correcting code (EPCC) is a code designed to correct frequently observed error cluster patterns of the intersymbol interference (ISI) channel. This paper focuses on developing theoretical understanding of the performance of serial concatenation of the EPCC with an outer recursive systematic convolutional code (RSCC) in ISI channel environments. To analyze the performance of this EPCC-RSCC concatenation, an upper union bound on the maximum-likelihood (ML) bit-error rate (BER), averaged over all possible interleavers, is derived which offers crucial insights into the error floor behavior of the matching turbo decoder. The ML bound is also used to compare the performance of EPCC-RSCC to that of a stand-alone RSCC in serial concatenation to precoded and nonprecoded ISI channels. This comparison shows that by targeting the low Hamming-weight interleaved errors of the RSCC, which result in low Euclidean distance error events in the channel detector, EPCC-RSCC exhibits a much lower BER floor compared to conventional schemes, especially for high rate applications and short interleaver lengths. The error rate performance of an iterative suboptimal turbo equalizer (TE), called TE-EPCC, is also demonstrated to converge close to the ML bound at high SNR.

Index Terms—Dominant error patterns, error pattern correcting code, error weight enumerator, intersymbol interference, list decoding, maximum-likelihood bit error rate bound, recursive systematic convolutional code, turbo equalization, uniform interleaver.

I. INTRODUCTION

TURBO CODES have become immensely popular in facilitating communication rates close to channel capacity. Soon after its inception [1], the turbo code principle was used to eliminate the effect of intersymbol interference (ISI) through the technique that has been called turbo equalization [2], [3]. Since

Manuscript received February 07, 2010; revised September 06, 2010; accepted November 09, 2010. Date of current version January 19, 2011. This work was supported in part by the NSF under Theoretical Foundation Grant 0728676 and IHCS Grant 0701946, and in part by the National Research Foundation of Korea under Grant 2010-0029205.

This paper is part of the special issue on “Facets of Coding Theory: From Algorithms to Networks,” dedicated to the scientific legacy of Ralf Koetter.

H. Alhussien was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with Link-A-Media Devices, Santa Clara, CA 95051 USA (e-mail:hakima@link-a-media.com).

J. Moon was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with the Department of Electrical Engineering, KAIST, Yuseong-gu, Daejeon, 305-701, Korea (e-mail:jmoon@kaist.edu).

Communicated by A. C. Singer, Associate Editor for the special issue on “Facets of Coding Theory: From Algorithms to Networks.”

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2010.2096031

then, the term turbo equalization has been used to refer to any technique that combines soft-in soft-out (SISO) decoding of an error correction code with SISO channel detection through iterative exchanging of soft information [3]–[5]. Codes that have been used in this setup include convolutional codes (CC), low-density-parity check (LDPC) codes, and turbo product codes (TPC). In this work we focus on turbo equalization based on the convolutional code.

The original turbo code is based on parallel concatenated convolutional codes (PCCCs) separated by an interleaver, for which the joint probability of two codes generating low Euclidean distance error events is shown to be considerably reduced by the action of the uniform interleaver [6]. This in effect improves the overall system bit-error-rate (BER) in the low-to-medium signal-to-noise ratio (SNR) region. The PCCCs are decoded by iterative exchange of soft information between the *a posteriori* probability (APP) processors matched to the constituent recursive systematic convolutional codes (RSCCs) [7]. A turbo equalizer composed of a PCCC soft decoder and a channel detector was discussed in [8]. A simpler serial concatenation of a single RSCC and a precoder through an interleaver was found to perform just as well in [9] for wireless communication applications, and in [10] and [11] for magnetic recording applications.

A precoder combined with the ISI channel behaves as a recursive rate-1 convolutional code [3], [12]. This combination resembles the serially concatenated convolutional codes (SCCCs), which were shown to perform better in terms of BER when the inner constituent code is recursive [13]. The function of a precoder is to reduce the frequency of low to medium Euclidean distance errors in the error weight spectrum of nonprecoded TEs [14], the phenomenon called “spectral thinning.” This results in a lower waterfall BER in the low to medium SNR region, but fails to improve the floor BER in the high SNR region. To improve the BER in both SNR regions, the concatenation of an inner cyclic code, called the the error-pattern correcting code (EPCC) [15], [16], and an outer RSCC through a bit interleaver, was recently considered in [17] for magnetic recording. Error pattern correction (EPC) coding is designed to correct the most probable error events at the output of the ISI-channel-matched sequence detector. Since EPCCs maintain a substantial error correction power while having a very high code rate, the hope is that the redistribution of redundancy between the EPCC and the outer RSCC would improve overall system performance.

The main contribution of this paper is to provide theoretical understanding of how the error weight spectrum is affected by the presence of the channel-matched EPCC in the EPCC-RSCC serial concatenation, and to develop necessary insights

into the proper design of the EPCC in this setup. We show how an EPCC can induce a drastic improvement in the frequency of low-weights in the error spectrum of the coded channel output, which leads to a much lower error floor compared to conventional TEs. To understand the error rate behavior of the EPCC-RSCC concatenation, we rely on the union bound approach assuming the maximum-likelihood (ML) decoder and adopting the well-established notion of the uniform interleaver [6]. The derivation of the ML bound presented in this paper, however, contains some elements unique to the study of the EPCC. While in previous methodologies all error events are treated the same, we distinguish error events in our derivation according to their contribution to the BER, and this distinction unveils the mechanism by which the EPCC achieves a dramatic spectral thinning without being a recursive inner code. Specifically, the Euclidean distance distribution modified by the EPCC is evaluated, and it is shown how the distribution is truncated for low Euclidean distances by the action of the EPCC for different multiple error correction capabilities. Furthermore, we provide single sum expressions of the ML BER as a function of the Hamming weight distribution of the outer RSCC, and in the case of more than one EPCC codewords per interleaver block, we provide an efficient evaluation method based on multinomial theory. Also, the parameter corresponding to the interleaver size raised to integer powers is isolated in the expressions of the derived ML bound, and this provides valuable insight into the interleaver gain of the EPCC as compared to that of the channel precoding.

Although interleavers function differently to achieve turbo gain in precoded and EPCC-based TEs, they are essential in the operation of both systems. In a precoded TE, the interleaver increases the average length of low Hamming-weight errors, and since longer errors result in larger average Euclidean distance errors, this improves the waterfall BER. In the EPCC-enhanced nonprecoded TE, the interleaver breaks large Hamming-weight single error events into multiple dabit error events with high probability. Then, since all multiple occurrences of the targeted dabit error events belong to the dominant error class, they are correctable by an EPCC of suitable multiple error correction capability.

The derivation of the upper bound on ML BER is based on a few assumptions that have already been employed in the related literature [6], [14], [18]. First, the bound is based on the notion of a uniform interleaver, which essentially averages out the effect of good and bad instantaneous interleavers on the bound. The implication of this assumption on the analytic BER bound is that the particular choice of the practical interleaver is not a factor in the turbo system comparison herein. Second, the derivation of the bound presumes a maximum-likelihood decoder, which falls short of accurately describing the iterative turbo gain that is more pronounced at lower to medium SNR. Note that the analysis of turbo code performance at this lower SNR region remains largely an open problem. Incidentally, our proposed approach here based on probabilistic correction of low Euclidean distance errors is designed to work in the floor region where the bound is accurate. Finally, the bound assumes that coded bits that go through the ISI channel are independent and identically distributed (i.i.d.), which becomes a more realistic approximation when the overall code rate is high [14].

Though the analysis here focuses on deriving an upper bound on the ML BER of the EPCC-RSCC concatenation, an ML decoder will be prohibitively complex to implement. Thus, to validate the ML bound analysis, we run bit error rate simulation of a suboptimal iterative decoder that converges close to the ML bound if carefully designed. In this suboptimal design, the EPCC decoder works iteratively with the outer RSCC decoder to correct low Euclidean distance errors at the output of the channel detector. This is compared to using the rate-1 precoder at the encoder side to prevent these errors from occurring in high frequency, but without eliminating them entirely.

The paper is organized as follows: In Section II we review the main concepts of EPC code construction and decoding based on its algebraic properties; we also present EPCC design examples that we later use in the simulation of Section VI. In Section III we summarize the encoder and suboptimal iterative decoder components of the conventional precoded and non precoded TEs (based on stand-alone RSCC) and of EPCC-RSCC concatenation. Section IV analyzes the ML BER performance of EPCC-RSCC and stand-alone RSCC based on the overall error weight spectrum of the coded channel. Furthermore, this section discusses an efficient method to evaluate the BER bound based on multinomial theory, assuming a single EPCC codeword per interleaver block. Section V highlights the main contribution of this paper: we explain the gains of the EPCC-RSCC over stand-alone RSCC in terms of the improved interleaver gain exponents of lower Euclidean distance errors. The numerical results presented in Section VI corroborate our claims in a variety of channel conditions for a combination of decoder design parameters. Finally, Section VII concludes the paper.

II. REVIEW OF THE ERROR-PATTERN-CORRECTING CODE

The cyclic codes described in [16] are based on construction of a generator polynomial $g(x)$ that gives rise to distinct syndrome sets for each targeted dominant error pattern. It has been shown that such a $g(x)$ can be obtained from the irreducible factors of the polynomial representations of the dominant error patterns. Furthermore, the code can be extended by introducing another factor in $g(x)$, namely, a primitive polynomial that is not already a factor of $g(x)$ [19]. This increases the code rate and codeword length, and also improves the single-error-pattern correction accuracy.

We start by constructing a cyclic code targeting the set of l_{\max} dominant error events

$$\{e^{(1)}(x), e^{(2)}(x), \dots, e^{(l_{\max})}(x)\}$$

represented as polynomials on $GF(2)$. An error event type i that starts at position $k \in \{0, 1, \dots, l_n - 1\}$ in a codeword of length l_n is given by $e_k^{(i)}(x) = x^k e^{(i)}(x)$. A syndrome of error $e^{(i)}(x)$ at position k is defined as $s_k^{(i)}(x) = e_k^{(i)}(x) \bmod g(x)$, where $g(x)$ is the generator polynomial of the code, and \bmod is the polynomial modulus operation. The syndrome set S_i for error $e^{(i)}(x)$ contains elements corresponding to all cyclic shifts of polynomial $e^{(i)}(x)$ in the codeword; the k th and j th elements of S_i are thus related by $s_j^{(i)} \equiv x^{j-k} s_k^{(i)} \bmod g(x)$. The period P_i of S_i is defined as the smallest integer such that $s_{\rho+P}^{(i)} = s_{\rho}^{(i)}$ [20].

For unambiguous decoding of $e^{(i)}(x)$ and $e^{(j)}(x), \forall\{i, j\}$, we must have $\mathbf{S}_i \cap \mathbf{S}_j = \emptyset$. This design requirement constrains $g(x)$ to have distinct greatest common divisors with all $e^{(i)}(x)$. However, even if this constraint is satisfied, an element in \mathbf{S}_i can still map to more than one position, i.e., the period of the syndrome set—and period of $g(x)$ —can be less than l_{\max} , where the period of $g(x)$ is equivalent to the polynomial order, which is the smallest n such that $g(x)$ divides $x^n + 1$. Moreover, this constraint is only sufficient but not necessary. Also, as shown in [16], there may exist a lower degree $g(x)$ that can yield distinct syndrome sets for the targeted error polynomials, resulting in a higher rate EPCC. A search method to find this $g(x)$ is already discussed in detail in [16] and [20].

We now describe the construction and properties of the EPCC that will be deployed throughout the paper. We target the dominant error events of a generalized two tap ISI channel of the form $1 - \alpha D, 0 < |\alpha| \leq 1$, for which the dicode and PR1 channels are special cases. Allowing $|\alpha|$ to deviate from 1 is useful for modelling practical recording channels [17]. When α is close to 1, the dominant errors are: $+, +-, +-$, etc., which have the polynomial representations: $e^{(1)}(x) = 1, e^{(2)}(x) = 1 + x, e^{(3)}(x) = 1 + x + x^2$, etc., i.e., polynomials on $GF(2)$ for which all powers of x have nonzero coefficients.

For the purpose of designing EPCCs for use in the TE-EPCC, the component EPCC code rate should be very high. To maintain a high rate, the EPCC codeword has to be extended in length without proportionally increasing the number of parity bits required to achieve accurate single-error occurrence correction capability. Example EPCCs are shown next, and the syndrome set periods of these codes are shown in Table I.

- (630, 616) EPCC: Targeting error polynomials up to degree 9, we get the generator polynomial $g(x) = 1 + x^3 + x^5 + x^8$ of period 30, via the search procedure in [16]. Choosing a codeword length of 30, 10 distinct, nonoverlapping syndrome sets are utilized to distinguish the 10 target errors. However, the resulting (30, 22) EPCC has rate 0.73 which incurs a high rate penalty. By multiplying the base EPCC generator polynomial by the primitive polynomial $1 + x + x^6$, which is not a factor of any of the targeted error polynomials, we obtain the extended generator polynomial $g_e(x) = 1 + x + x^3 + x^4 + x^5 + x^8 + x^{11} + x^{14}$, which corresponds to the extended (630, 616) EPCC of rate 0.98, and 14 parity bits. Then, as shown in [16], syndrome sets $\mathbf{S}_1, \mathbf{S}_3, \mathbf{S}_7$, and \mathbf{S}_9 have period 630 and thus can be decoded without ambiguity. On the other hand, syndrome sets $\mathbf{S}_2, \mathbf{S}_4, \mathbf{S}_6$, and \mathbf{S}_8 have period 315, decoding to one of two positions. The worst would be \mathbf{S}_5 of period 126, and \mathbf{S}_{10} of period 63, which decode to 5 and 10 possible positions, respectively. Still, the algebraic decoder can quickly shrink the number of possible error positions to a very few positions by checking the data support, and then would choose the one position with highest local reliability, which is calculated using the channel observations and *a priori* side information.
- Shortened (126, 112) EPCC: Shorter lower-rate EPCCs can be obtained by shortening the (630, 616) EPCC. For example, a (126, 112) EPCC of rate 0.89 can be derived this

TABLE I
SYNDROME SET PERIODS OF VARIOUS EPCCS

Target error	(630, 616) EPCC	(126, 112) EPCC	(210, 199) EPCC
1	630	126	210
$(1 + x)$	315	126	105
$(1 + x + x^2)$	630	126	70
$(1 + x)^3$	315	126	105
$(1 + x + x^2 + x^3 + x^4)$	126	126	42
$(1 + x)(1 + x + x^2)^2$	315	126	35
$(1 + x + x^3)(1 + x^2 + x^3)$	630	126	—
$(1 + x)^7$	315	126	105
$(1 + x + x^2)(1 + x^3 + x^6)$	630	126	70
$(1 + x)(1 + x + x^2 + x^3 + x^4)^2$	63	63	21

way with all syndromes sets, excluding syndrome set \mathbf{S}_{10} , having period 126, and thus are decodable without ambiguity.

- (210, 199) EPCC: To obtain short EPCCs without jeopardizing the code rate through code shortening, we can target fewer error patterns in the code design. Targeting error polynomials up to degree 9, but excluding $e^{(7)}(x)$, we can extend the base generator polynomial $g(x) = 1 + x^3 + x^5 + x^8$ through its multiplication by the primitive polynomial $1 + x + x^3$, which we could not use before because it is a factor of the polynomial representation of $e^{(7)}(x)$. The resulting code is a (210, 199) EPCC of rate 0.95, 11 parity bits, and extended generator polynomial $g_e(x) = 1 + x + x^4 + x^5 + x^9 + x^{11}$.

III. RSCC AND EPCC CONCATENATION SCHEME

The concatenation of RSCC and the nonprecoded ISI channel is shown in Fig. 1(i). In the encoder side, a feedback shift register encodes the data stream, which is interleaved before being passed to the channel. The concatenation of the convolutional code and the ISI channel can be viewed in the context of turbo coding as a serial concatenation of an outer recursive code and an inner rate-1 nonrecursive code through an interleaver. The one-shot ML receiver for this concatenation is prohibitive in complexity. However, a suboptimal polynomial-time iterative-type decoder can be designed based on the separation of detection and decoding. The separate APP decoder and APP detector iteratively exchange soft bit decisions, with the error rate performance converging close to that of the combined ML solution at high SNR [13]. APP processing can be realized via the soft output Viterbi algorithm (SOVA) [21]. A suboptimal iterative receiver can also be built using other APP processors, such as the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [22], or the minimum mean squared error (MMSE) soft-in soft-out (SISO) detector of [4], [5].

The BER gain in turbo equalization is most notable at low SNRs, and plateaus rapidly as SNR increases resulting eventually in the error floor phenomenon. The gain at low SNR is further enhanced by including a rate-1 recursive component in the

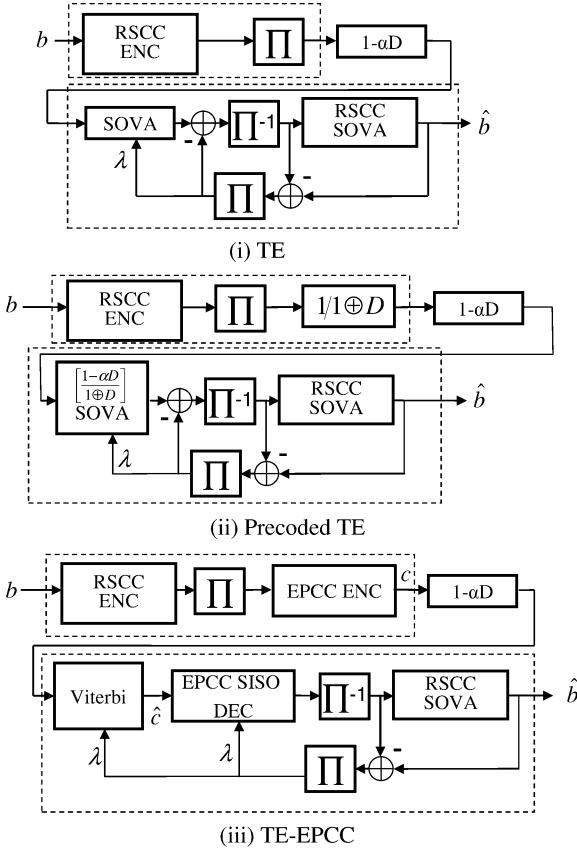


Fig. 1. Block diagrams of: (i) the TE matched to the stand-alone RSCC without precoding; (ii) the TE matched to the stand-alone RSCC with precoding; and (iii) the TE-EPCC matched to the EPCC-RSCC concatenation.

path of the RSCC coded interleaved bit stream. This is shown in Fig. 1(ii), and is called the precoded TE. The suboptimal iterative receiver is modified to account for precoding, where the trellis of the SOVA is now matched to the recursive rate-1 coded channel $\frac{1-\alpha D}{1+\alpha D}$. By the action of the uniform interleaver, the fraction of errors in the Hamming-weight error distribution of the RSCC resulting in low Euclidean distance errors in the channel trellis is greatly reduced. This, as a result, improves the BER at low to medium SNRs, wherein the contribution of the reduced profile of error Euclidean distances to the BER exceeds the minimum distance effect.

A markedly different approach is depicted in Fig. 1(iii). This method is based on replacing the rate-1 precoder with a very high rate EPCC that is designed to correct low Hamming-weight errors that generate low Euclidean distance trellis errors. EPCC coding shifts the paradigm from constraining the occurrence of low Euclidean distance errors to postdetection correction of a large fraction of these errors.

While the $\frac{1}{1+\alpha D}$ precoder and the $1 - \alpha D$ channel can be jointly decoded/detected with no added complexity by matching the trellis to the combined coded channel $\frac{1-\alpha D}{1+\alpha D}$, it is impractical to realize a similar one-shot ML decoder of the channel and the EPCC. Hence, in Fig. 1(iii) separate decoders of the channel and the EPCC are implemented to construct the suboptimal iterative receiver, and since the EPCC is matched to the ISI channel, no interleaving should be present between its encoder (or decoder)

and the channel. On the other hand, an interleaver is essential between the EPCC and the outer RSCC. A SISO decoder for the EPCC is discussed in [20]. An EPCC SISO decoder that runs in a TE setting is also discussed in [17].

IV. ML ERROR-RATE ANALYSIS OF EPCC-RSCC CONCATENATION

In bounding the BER of EPCC-RSCC, we closely follow the basic approach of [14], [23] to condition the probability of Euclidean distance at the ISI channel output on the Hamming distance between a pair of interleaved codewords. Our analysis deviates from that of [14], [23] where we discuss classification of channel error patterns into dominant versus nondominant classes. Also, to cover practical recording channels, we apply the analysis to a slightly generalized two-tap channel of the form $1 \pm \alpha D$. The dicode ($1 - D$) and PR1 ($1 + D$) channels are special cases corresponding to $\alpha = 1$. Given the BER established as a function of the error Euclidean distance distribution of the overall system, we argue for EPCC-RSCC's enhanced performance by the virtue of its ability to reduce occurrence frequencies of low Euclidean distances.

Following the notations of [14], the ML union bound on the word error rate of a block code of codebook size M , of equally likely codewords, and AWGN of zero mean and variance σ^2 is

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

where m and \tilde{m} point to codewords separated by Euclidean distance $\|\mathbf{x}_m - \mathbf{x}_{\tilde{m}}\|$, and \mathbf{x}_m is the noiseless channel output corresponding to m . If there are T_{m,d_E} different codewords for which the corresponding noiseless channel outputs are at distance d_E from \mathbf{x}_m , then we can write (1) as

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

where $\bar{T}(d_E)$ is the average number of codewords at Euclidean distance d_E from a codeword, with the distance measured at the channel output. The associated BER can be shown to be

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

where K is the number of information bits per codeword, and $\bar{w}(d_E)$ is the average Hamming distance measured from an information word to competing information words located at d_E away, with the Euclidean distance measure based on noiseless channel outputs of the corresponding codewords. As in [14], the basic approach is to relate the function $\bar{T}(d_E)$ to the outer code Hamming weight enumerator $A^O(d)$ and the error event characteristics of the channel, but in our case the effect of the inner

EPCC also comes into play through a modified $\bar{T}(d_E)$ specialized to the set of channel error events that are not correctable by the EPCC.

A. Error Event Analysis of the $1 - \alpha D$ Channel

A trellis section of the $1 - \alpha D$ channel with no precoding is shown in Fig. 2. The branch label c_i/x_i signifies the coded input bit to the channel and the corresponding channel output, respectively. Again, using the same notations as in [14], any error word f with Hamming weight $d = d^H(f)$ can be uniquely decomposed into disjoint error patterns (or error subevents) $f_j, j = 1, \dots, m$, where the index j signifies the position of the error pattern of Hamming weight d_j^H within the codeword block. Error patterns $f_j, j < m$, correspond to simple closed error events on the trellis that diverge from and remerge into the correct path without sharing any of the states in between. The last subevent f_m either remerges with the correct path (closed f_m) or remains diverged even after the boundary of the codeword has been reached (open f_m) [14]. In the $1 - \alpha D$ channel trellis, diverging branches result in a Euclidean distance separation of 1 each, while remerging branches result in a squared Euclidean distance separation of α^2 each. Moreover, crossing branches accumulate a squared distance separation of $(1 + \alpha)^2$, while parallel branches accumulate a separation of $(1 - \alpha)^2$. This means that parallel branches result in a lower Euclidean distance separation compared to crossing branches in the Euclidean distance distribution when $0 < \alpha \leq 1$. Hence, two error pattern classes are distinguishable according to their accumulated Euclidean distance. The first class, shown in Fig. 2(b), has a squared distance $d_E^2(f_j) = 1 + (1 - \mu)\alpha^2 + (d_j^H - 1) \times (1 - \alpha)^2$ where d_j^H is the Hamming weight of the subevent f_j , and $\mu = 1$ or 0 depending on the event being open or closed, respectively. This class of error patterns is denoted by χ^{dom} and is called the “dominant error class,” for which all branches, except for the diverging and remerging ones, are parallel.

The dominant error class accounts for most of the channel bit errors due to the low Euclidean distance between the correct and erroneous paths. On the other hand, the second class, shown in Fig. 2(c), has both parallel and crossing branching, and hence its members have squared Euclidean distance $d_E^2(f_j) = 1 + (1 - \mu)\alpha^2 + \lambda_j^{\text{cr}} \times (1 + \alpha)^2 + (d_j^H - 1 - \lambda_j^{\text{cr}}) \times (1 - \alpha)^2$, where λ_j^{cr} is the number of crossing branches. The second class contributes much less to the overall system BER, and thus we call it the “nondominant error class,” which is denoted by $\tilde{\chi}^{\text{dom}}$. By the same line of reasoning, the two classes can also be defined for the PR1 channel, a special case of $1 + \alpha D$ at $\alpha = 1$. The only difference is that error events with all crossing branches

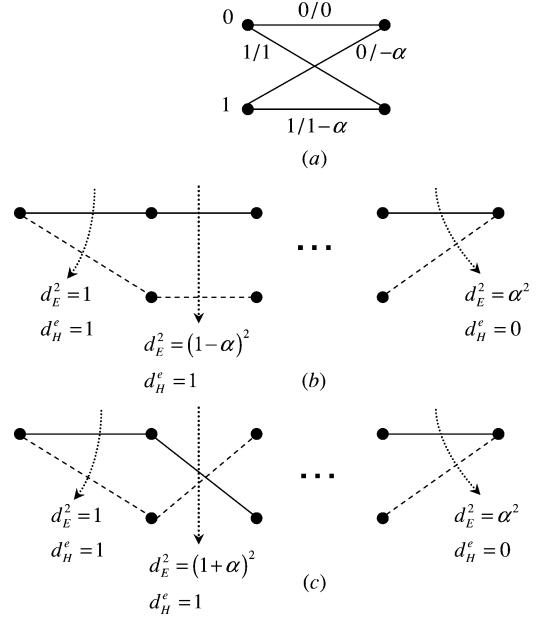


Fig. 2. (a) Trellis section for a nonprecoded two-tap ISI channel ($1 - \alpha D$). (b) A dominant error pattern. (c) A nondominant error pattern.

now generate the class χ^{dom} . The EPCC is capable of correcting error words f that are decomposable into disjoint error patterns f_j that all belong to the dominant error class, i.e., $f_j \in \chi^{\text{dom}}, \forall j$. In order to evaluate the BER performance of EPC coding, we need to find the new Euclidean distance distribution modified by the EPCC. However, it would be easier to first find the Euclidean distance distribution before EPC is turned on, a case leading to a simple generalization of the results given in [14]. We assume throughout that the code bit values at the channel input are i.i.d. and equiprobable within each error subevent, which is a valid assumption for a high rate code and is necessary to make statistical characterization of the Euclidean distance distribution at the channel output independent of the code constraints [11], [14]. In an error word f of Hamming weight $d^H(f) = d$, a dominant error pattern $f_j \in \chi^{\text{dom}}$ of length $l_j = d_H(f_j)$ contains no crossing branches and occurs with probability $(\frac{1}{2})^{l_j-1}$. On the other hand, a nondominant error pattern $f_j \in \tilde{\chi}^{\text{dom}}$ of length l_j and λ_j^{cr} crossing branches will have probability $\binom{l_j-1}{\lambda_j^{\text{cr}}} (\frac{1}{2})^{l_j-1}$. Therefore, the probability distribution of $d_E^2(f)$ is given by (4), at the bottom of the page, which is the conditional probability of an error word of Euclidean distance d_E (at the channel output), given that its Hamming weight

$$\Pr(d_E|d, m, \mu) = \begin{cases} \binom{d-m}{\lambda^{\text{cr}}} \left(\frac{1}{2}\right)^{d-m}, & \lambda^{\text{cr}} > 0 \text{ integer} \\ \left(\frac{1}{2}\right)^{d-m}, & \lambda^{\text{cr}} = 0 \\ 0, & \text{otherwise.} \end{cases}$$

$$\lambda^{\text{cr}} = \frac{d_E^2 - (1 - \alpha)^2 d - 2\alpha m + \mu\alpha^2}{4\alpha}, \quad (4)$$

is d and it has m error pattern occurrences (at the interleaver output), μ boundary errors, and λ^{cr} crossing branches.

For the precoded $1 - \alpha D$ trellis in Fig. 3, we note that a bit error results in the divergence of a single error event that remerges only on the occurrence of another bit error, while all the trellis branches in between have zero Hamming weight separation, whether crossing or parallel. We also note that an error event with even d^H decomposes into $\frac{d^H}{2}$ closed single error subevents, while that with an odd d^H decomposes into $\lfloor \frac{d^H}{2} \rfloor$ closed subevents and a single open subevent. Moreover, diverging and remerging branches have $d_E^2 = 1$ and $d_E^2 = \alpha^2$, respectively, while parallel and crossing branches have $d_E^2 = (1 - \alpha)^2$ and $d_E^2 = (1 + \alpha)^2$, respectively. This means that, by invoking the uniform interleaver assumption, the probability of a single long error event of $d^H \in \{1, 2\}$ producing a low average Euclidean distance error declines exponentially in the interleaver size, since the probability of an all parallel error event declines accordingly. The Euclidean distance associated with an error word with multiple subevents of total Hamming weight d , λ^{cr} crossing branches, and total length L is

$$d_E^2 = \left\lceil \frac{d}{2} \right\rceil + \left\lceil \frac{d}{2} \right\rceil \alpha^2 + (1 + \alpha)^2 \lambda^{\text{cr}} + (1 - \alpha)^2 (L - d - \lambda^{\text{cr}})$$

leading to

$$\Pr(d_E|d, L) = \begin{cases} \binom{L-d}{\lambda^{\text{cr}}} \left(\frac{1}{2}\right)^{L-d}, & \lambda^{\text{cr}} > 0 \text{ integer.} \\ 0, & \text{otherwise.} \end{cases}$$

$$\lambda^{\text{cr}} = \frac{d_E^2 - \lceil \frac{d}{2} \rceil - \lfloor \frac{d}{2} \rfloor \alpha^2 - (1 - \alpha)^2 (L - d)}{4\alpha}. \quad (5)$$

B. Error Euclidean Distance Distribution of EPCC-RSCC

We now develop a method to construct the Euclidean distance distribution of the error words for EPCC-RSCC as applied to the ISI channel. Consider a serial concatenation of an EPCC and a RSCC of length N . There are L_c EPCC

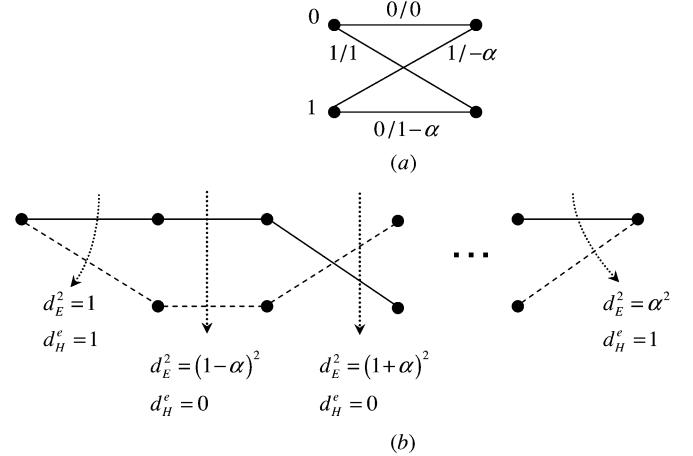


Fig. 3. (a) Trellis section for a $\frac{1}{1+D}$ precoded two-tap ISI channel ($1 - \alpha D$). (b) Weight characterization of an error pattern.

inner codewords in each interleaved block, each of user-length $N_c = \frac{N}{L_c}$. The EPCC is assumed to be capable of correcting up to m_c occurrences per codeword provided that they all belong to the target set of correctable errors. The target set is $\{\mathbf{f}_j : \mathbf{f}_j \in \chi^{\text{dom}}, d^H(\mathbf{f}_j) \leq d_c\}$, where d_c is the maximum length of correctable error patterns in χ^{dom} . The errors in a RSCC error codeword of Hamming weight d are mapped by the uniform interleaver into all possible $\binom{N}{d}$ interleaved error words \mathbf{f} with equal probability. The interleaved error word divides into L_c EPCC codewords, each receiving error word $\mathbf{f}^{(i)}, i = 1, \dots, L_c$, of Hamming weights d_1, d_2, \dots, d_{L_c} . Each EPCC error word $\mathbf{f}^{(i)}$ of Hamming weight d_i decomposes into m_i disjoint error pattern occurrences. In the previous section, we found the conditional probability $P(d_E|d, m, \mu)$ given the error Hamming weight and the number of errors m for a single codeword per interleaver block. To derive the Euclidean distance distribution for a codeword that is divisible into L_c codewords, we are also required to evaluate the conditional probability of the decompositions m_i given the EPCC codeword Hamming weights d_i . The conditional Euclidean distance probability distribution can be expanded as shown in (6), at the bottom of the page. Since errors in the L_c EPCC codewords are disjoint

$$\begin{aligned} \Pr(d_E|d) &= \sum_{d_1=0}^d \cdots \sum_{d_{L_c}=0}^d \Pr(d_E|d, d_1, \dots, d_{L_c}) \\ &\quad \times \Pr(d_1, \dots, d_{L_c}|d) \\ &= \sum_{d_1=0}^d \cdots \sum_{d_{L_c}=0}^d \sum_{m_1=0}^{d_1} \cdots \sum_{m_{L_c}=0}^{d_{L_c}} \Pr(d_E|d, d_1, \dots, d_{L_c}, m, m_1, \dots, m_{L_c}) \\ &\quad \times \Pr(m, m_1, \dots, m_{L_c}|d_1, \dots, d_{L_c}, d) \\ &\quad \times \Pr(d_1, \dots, d_{L_c}|d). \end{aligned} \quad (6)$$

(i.e., $\Pr(m, m_1, \dots, m_{L_c} | d_1, \dots, d_{L_c}, d) = \prod_{i=1}^{L_c} \Pr(m_i | d_i)$), (6) becomes

$$\begin{aligned} \Pr(d_E | d) &= \sum_{d_1=0}^d \dots \sum_{d_{L_c}=0}^d \Pr(d_1, \dots, d_{L_c} | d) \\ &\quad \times \sum_{m=1}^d \sum_{m_1=0}^{d_1} \dots \sum_{m_{L_c}=0}^{d_{L_c}} \Pr(d_E | d, m) \prod_{i=1}^{L_c} \Pr(m_i | d_i) \quad (7) \end{aligned}$$

The joint conditional probability $\Pr(d_1, \dots, d_{L_c} | d)$ in (7) is the probability of the interleaved error word \mathbf{f} , of Hamming weight d , mapping into the error word sequence $\{\mathbf{f}^{(i)}\}_{i=1}^{L_c}$ with associated Hamming weight sequence $\{d_i\}_{i=1}^{L_c}$ out of all possible $\binom{N}{d}$ interleaved words, and is given by

$$\Pr(d_1, \dots, d_{L_c} | d) = \frac{\binom{N_c}{d_1} \times \binom{N_c}{d_2} \dots \times \binom{N_c}{d_{L_c}}}{\binom{N}{d}}. \quad (8)$$

Given that there are d_i errors in EPCC codeword i , there exists $\binom{d_i - 1}{m_i - 1}$ ways by which the d_i errors are distributed into m_i error pattern occurrences, each of length at least 1. Of these m_i occurrences, $\mathbf{f}_{m_i}^{(i)}$ can be either open or closed. An open error event in this context lies on the boundary of the EPCC codeword's data and parity fields. We will ignore the Euclidean distance growth due to errors in the parity fields, which will result in a slightly pessimistic error rate expression for the proposed EPCC-RSCC scheme, but this will allow us to proceed with the analysis. We also assume a fixed initial state for all EPCC codewords.

By examining the trellis we note that a boundary error subevent contributes to the squared Euclidean distance separation by α^2 less than a closed subevent with identical length. Furthermore, there are only $\binom{N_c - d_i}{m_i - 1}$ ways by which the disjoint m_i error patterns of error word $\mathbf{f}^{(i)}$ can be arranged in the current codeword i , given the codeword has a boundary error. Note that two disjoint error occurrences in the trellis are separated by at least the error free distance of the channel, which equals 1 for $1 \mp \alpha D$ ISI channels, and the number of possible arrangements of m_i errors is computed given the fact that the last error pattern occurs at the boundary. On the other hand, assuming errors can occur on and off the boundary, the total number of possible error pattern arrangements becomes $\binom{N_c - d_i + 1}{m_i}$. We will further assume that open error subevents have very low probability of extending into the data fields of adjacent EPCC codewords, and hence, error events are independent among different codewords.

So, given μ_i open subevents per error word i , there are $\binom{N_c - d_i}{m_i - \mu_i}$ ways by which the m_i error patterns, composing

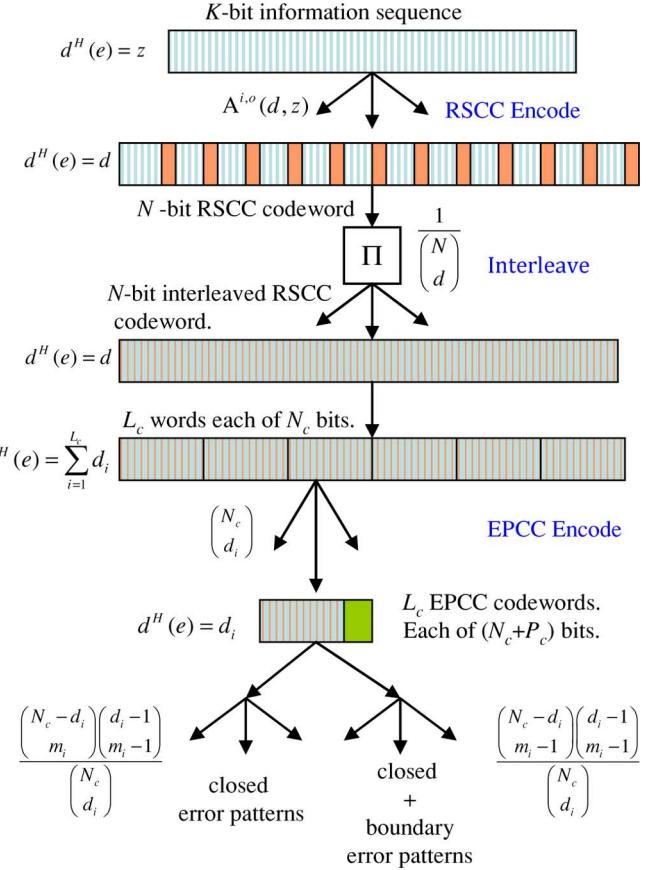


Fig. 4. Sketch of the method to derive $\bar{T}(d_E)$.

$\mathbf{f}^{(i)}$, can be arranged in codeword i . Since there are $\binom{N_c}{d_i}$ possible error words $\mathbf{f}^{(i)}$, we get

$$\Pr(m_i | \mu_i, d_i) = \frac{\binom{N_c - d_i}{m_i - \mu_i} \times \binom{d_i - 1}{m_i - 1}}{\binom{N_c}{d_i}}. \quad (9)$$

A pictorial depiction of the derivation method explained above is shown in Fig. 4. Substituting (4), (9), and (8) into (7), we get an expression for the distribution of error Euclidean distances while EPC is turned off as

$$\begin{aligned} \Pr(d_E | d) &= \sum_{m=1}^d \sum_{\mu=0}^{L_c} \sum_{i=1}^{L_c} \Phi(m, \mu, \alpha | d) \\ &\quad \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{\alpha} = 0 \bmod 4 \\ \Phi(m, \mu, \alpha | d) &= \sum_{d_i=0}^d \sum_{m_i=0}^{d_i} \sum_{\mu_i=0}^1 \xi^{(i)}(\alpha) \\ &\quad d_i = \sum_i d_i, m = \sum_i m_i, \mu = \sum_i \mu_i \\ \xi^{(i)}(\alpha) &= \frac{1}{\binom{N}{d}} \left(\frac{d - m}{\frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{4\alpha}} \right) \left(\frac{1}{2} \right)^{d-m} \\ &\quad \times \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} \end{aligned} \quad (10)$$

TABLE II
INTERLEAVER GAIN OF RSCC WITH NO PRECODING VERSUS EPCC-RSCC, $d_E^2 = \{1, 2, 3, 4\}$

$d_E^2 = 1$	Error pattern classes	RSCC	EPCC-RSCC
$m = 1$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-2} (1)	N^{-11} ($\frac{155925}{4}$)
$d_E^2 = 2$	Error pattern classes	RSCC	EPCC-RSCC
$m = 1$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^{-1} (1)	N^{-10} ($\frac{155925}{4}$)
$d_E^2 = 3$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-1} (2)	N^{-10} (779625)
$d_E^2 = 4$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^0 (1)	N^{-9} ($\frac{779625}{2}$)

where we define $\binom{0}{0} = 1$. In addition, the Euclidean distance distribution can be decomposed into two components: a component $\Pr(d_E, \mathcal{C}|d)$ associated with RSCC error words that are correctable by the L_c EPCC codewords, and the complimentary component $\Pr(d_E, \tilde{\mathcal{C}}|d)$ associated with noncorrectable RSCC error words. In this case, the Euclidean distance probability distribution of noncorrectable error words escaping EPCC-RSCC is given by

$$\Pr(d_E, \tilde{\mathcal{C}}|d) = \Pr(d_E|d) - \Pr(d_E, \mathcal{C}|d) \quad (11)$$

while the correctable component is given by

$$\begin{aligned} \Pr(d_E, \mathcal{C}|d) &= \sum_{m=1}^d \sum_{\mu=1}^{L_c} \prod_{i=1}^{L_c} \Phi(m, \mu, \alpha|d) \\ \Phi(m, \mu, \alpha|d) &= \sum_{\substack{d_i=0 \\ d=\sum_{i=1}^{L_c} d_i, m=\sum_{i=1}^{L_c} m_i, \mu=\sum_{i=1}^{L_c} \mu_i}} \sum_{m_i=0}^{\min(d_i, d_c)} \sum_{\mu_i=0}^{\min(d_i, m_i)} \xi^{(i)}(\alpha) \\ \xi^{(i)}(\alpha) &= \frac{1}{\binom{N}{d}} \left(\frac{1}{2}\right)^{d-m} \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} \end{aligned} \quad (12)$$

where for the sake of simplicity, we assumed that EPCC codeword i could correct an error word $\mathbf{f}^{(i)}$ if $d_H(\mathbf{f}^{(i)}) \leq d_c$, which is actually a worst case scenario that occurs only if $m_i = 1$. (Note that the EPCC decoder in practice can correct multiple error patterns whose combined Hamming weight could be more than d_c .) Although this assumption would reduce the cardinality of the set of EPCC-correctable errors and loosen the upper bound, it allows us to avoid a substantially more complicated derivation.

To obtain the bound on the bit error probability, we need to express the error Euclidean distance enumerators as a function of the error Euclidean distance probability distribution given by (11). We note that the average Euclidean distance enumerator associated with the uncorrectable set of error words $\tilde{\mathcal{C}}$ is given by

$$\bar{T}(d_E, \tilde{\mathcal{C}}) = \sum_{d=d_{\min}}^N A^O(d) \Pr(d_E, \tilde{\mathcal{C}}|d) \quad (13)$$

while the average input Hamming distance to codewords at squared Euclidean distance d_E^2 is given by

$$\bar{w}(d_E, \tilde{\mathcal{C}}) = \frac{1}{\bar{T}(d_E, \tilde{\mathcal{C}})} \sum_{d=d_{\min}}^N A^O(d) \bar{A}^I(d) \Pr(d_E, \tilde{\mathcal{C}}|d) \quad (14)$$

where $A^O(d)$ represents the number of RSCC codewords of weight d , and $\bar{A}^I(d)$ represents the average input Hamming weight of RSCC codewords of weight d , and are related by

$$\bar{A}^I(d) = \frac{\sum_z z A^{I,O}(d, z)}{A^O(d)}. \quad (15)$$

where $A^{I,O}(d, z)$ is the number of codewords of weight d that originated from weight z information words. Details on how to find these marginal error weight enumerators can be found in [24] for all the puncturing rates and encoder connection polynomials that are used in this paper. By substituting $\bar{T}(d_E, \tilde{\mathcal{C}})$, given by (13), and $\bar{w}(d_E, \tilde{\mathcal{C}})$, given by (14), in (3), we get an upper bound on the average BER of EPCC-RSCC as function of $\Pr(d_E|d, \tilde{\mathcal{C}})$

$$P_b \leq \sum_{d_E=d_{\min}}^{\infty} \sum_{d=d_{\min}}^N \frac{A^O(d) \bar{A}^I(d) \Pr(d_E, \tilde{\mathcal{C}}|d)}{K} Q\left(\frac{d_E}{2\sigma}\right). \quad (16)$$

TABLE III
INTERLEAVER GAIN OF RSCC WITH NO PRECODING VERSUS EPCC-RSCC, $d_E^2 = \{5, 6\}$

$d_E^2 = 5$	Error pattern classes	RSCC	EPCC-RSCC
$m = 3$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-1} (3)	N^{-9} (3508313)
$d_E^2 = 5$	Error pattern classes	RSCC	EPCC-RSCC
$m = 1$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-2} (1)	N^{-2} (1)
$d_E^2 = 6$	Error pattern classes	RSCC	EPCC-RSCC
$m = 3$ $\mu = 0$ $d = 3 \rightarrow d_T$		N^0 (1)	N^{-8} (1169438)
$d_E^2 = 6$	Error pattern classes	RSCC	EPCC-RSCC
$m = 1$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^{-1} (1)	N^{-1} (1)

In Appendix A we show how these bounds simplify in the simple case when $L_c = 1$, i.e., employing one EPCC codeword per interleaver block. Also, we extend the BER bound derived in [14] for the precoded dicode channel to the generalized case $1 - \alpha D$. Finally, by using an exponential-type approximation of the Q function, we show in Appendix A that the ML BER bounds of the compared systems can be expressed as single infinite sums over the Hamming weight of the RSCC error word.

C. Efficient Computation of the Euclidean Distance Enumerator for $L_c > 1$ EPCC

A more compact and efficient method is derived here to evaluate the multiple summations in (12) and (10), which are used to compute the BER bound in (16). We first define a probability enumerator for codeword i for all possible values of the parameters d_i, m_i and μ_i , which is given by the multinomial

$$\begin{aligned} \Lambda(D, M, \Upsilon; m_{\max}, d_{\max}) \\ = 1 + \sum_{\mu_i=0}^1 \sum_{d_i=1}^{d_{\max}} \sum_{m_i=1}^{\min(d_i, m_{\max})} \left(\frac{1}{2}\right)^{d_i-m_i} \\ \times \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} D^{d_i} M^{m_i} \Upsilon^{\mu_i} \end{aligned} \quad (17)$$

where the $D^0 M^0 \Upsilon^0 = 1$ monomial term corresponds to the case when there are no errors in the specified codeword, and $\mu_i = \{0, 1\}$ is the number of boundary subevents per codeword. As a result, the probability enumerator for the entire interleaver block composed of L_c EPCC codewords is given by

$$\Lambda^{L_c}(D, M, \Upsilon; m_{\max}, d_{\max})$$

given that only the weight- d_{\max} error words $f^{(i)}$ composed of m_{\max} disjoint error patterns can occur per EPCC codeword,

where d_{\max} and m_{\max} are unbounded from above if EPC is turned off. The advantage of this approach is that polynomial multiplication, or the more general multinomial multiplication, can be performed efficiently by symbolic manipulators, such as Maple, speeding up the evaluation of (12) and (10). Utilizing the compact, and easy-to-compute, probability enumerator, we can now express the bound on the bit error rate of the EPCC-RSCC as

$$\begin{aligned} P_b \leq \frac{1}{K} \sum_{d_E=d_E^{\min}}^{\infty} Q\left(\frac{d_E}{2\sigma}\right) \sum_{d=d_{\min}}^N \frac{A^O(d) \bar{A}^I(d)}{\binom{N}{d}} \\ \times \sum_{\mu=0}^{L_c} \sum_{\substack{m=1 \\ m: \lambda^{\text{cr}} \geq 0, \lambda^{\text{cr}} \in \mathbb{N}}}^d \binom{d-m}{\lambda^{\text{cr}}} \\ \times [\Lambda^{L_c}(D, M, \Upsilon; \infty, \infty)]_{d, m, \mu} \\ - \sum_{\substack{m=1 \\ m: \lambda^{\text{cr}}=0}}^d [\Lambda^{L_c}(D, M, \Upsilon; m_c, d_c)]_{d, m, \mu} \\ \lambda^{\text{cr}} = \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{4\alpha} \end{aligned} \quad (18)$$

where the probability enumerator for L_c correctable EPCC codewords is approximated by

$$\Lambda^{L_c}(D, M, \Upsilon; m_c, d_c)$$

for an EPCC of maximum correction power m_c per codeword, and \mathbb{N} is the set of natural numbers.

V. INTERLEAVER GAIN EXPONENT OF EPCC-RSCC

To gain insight into how EPC coding enhances TE performance in the asymptote of large interleaver size, we derive simpler analytic expressions of the bounds above that are direct

functions of the interleaver size. For this, we limit our investigation to the dicode channel, for which the spectrum of the Euclidean distance is comprised only of integer values of d_E^2 , and hence there are a few values that d_E can take in the lower range of the spectrum. In turbo coding, the coefficients of the error function for low Euclidean distances are an inverse function of the interleaver size, N , leading to what is widely known as the interleaver gain.

Below, we will argue for the advantage of incorporating an EPCC as an inner code, by showing how it works to rapidly decrease the exponent of N^δ well below zero, especially for low Euclidean distance errors. First, we isolate the term N^δ in the expression of ML BER for stand-alone RSCC and EPCC-RSCC. The ML BER expression of the stand-alone RSCC for $\alpha = 1$ is

$$\begin{aligned} P_b \leq \frac{1}{K} \sum_{d_E=1}^{\infty} Q\left(\frac{d_E}{2\sigma}\right) \sum_{d=2}^{d_T} \frac{A^O(d)\bar{A}^I(d)}{\binom{N}{d}} \\ \times \sum_{\mu=0}^1 \sum_{\substack{m=1 \\ m: d_E^2 - 2m + \mu = 0 \bmod 4}}^d \binom{d-m}{\frac{d_E^2 - 2m + \mu}{4}} \\ \times \left(\frac{1}{2}\right)^{d-m} \binom{N-d}{m-\mu} \binom{d-1}{m-1} \end{aligned} \quad (19)$$

where $d_T \ll N$ is the maximum Hamming-weight of error words in the truncated RSCC distribution, where the truncation is done such that thrown-out weights have little contribution to the BER, as they will be associated with large negative δ 's. To produce an expression for the upper bound on BER with isolated powers of N while preserving the bound as an upper bound, we replace the binomial in the denominator by the lower bound [13]

$$\binom{N}{d} > \frac{(N-d+1)^d}{d!} \simeq \frac{N^d}{d!}.$$

Moreover, to replace the binomial in the numerator with an upper bound that is also a power of $N-d+1$, we first express it as

$$\binom{N-d}{m-\mu} = \frac{m-\mu+1}{N-d+1} \binom{N-d+1}{m-\mu+1}$$

and employ the upper bound [13]

$$\binom{N-d+1}{m-\mu+1} < \frac{(N-d+1)^{m-\mu+1}}{(m-\mu+1)!} \simeq \frac{N^{m-\mu+1}}{(m-\mu+1)!}.$$

These bounds are tight when N is large and $d, m \ll N$, which holds true in our case. Also we can upper bound the Q function by $Q\left(\frac{d_E}{2\sigma}\right) \leq \frac{1}{2} e^{-\frac{d_E^2}{8\sigma^2}}$. Substituting these approximate bounds in the BER upper bound in (19), we get a looser but insightful bound

$$P_b < \frac{1}{2K} \sum_{d_E=1}^{\infty} \sum_{d=2}^{d_T} \sum_{\mu=0}^1 \sum_{\substack{m=1 \\ m: d_E^2 - 2m + \mu = 0 \bmod 4}}^d \Theta_{d_E, d, m, \mu} N^{m-\mu-d} e^{-\frac{d_E^2}{8\sigma^2}} \quad (20)$$

where $\Theta_{d_E, d, m, \mu}$ is given by

$$\begin{aligned} \Theta_{d_E, d, m, \mu} = A^O(d)\bar{A}^I(d) \frac{d!}{(m-\mu)!} \\ \times \left(\frac{1}{2}\right)^{d-m} \binom{d-m}{\frac{d_E^2 - 2m + \mu}{4}} \binom{d-1}{m-1} \end{aligned} \quad (21)$$

For the sake of mathematical tractability, we study the interleaver gain exponent of $L_c = 1$ EPCC-RSCC, i.e., single EPCC codeword per interleaver block. Utilizing the same approximations as above in the BER bound of EPCC-RSCC for $L_c = 1$ we get the expression

$$\begin{aligned} P_b < \frac{1}{2K} \sum_{d_E=1}^{\infty} e^{-\frac{d_E^2}{8\sigma^2}} \\ \times \sum_{\mu=0}^1 \left[\sum_{d=2}^{d_T} \sum_{\substack{m=1 \\ m: d_E^2 - 2m + \mu = 0 \bmod 4}}^d \Theta_{d_E, d, m, \mu} N^{m-\mu-d} \right. \\ \left. - \sum_{d=2}^{\min(d_T, d_c)} \sum_{\substack{m=1 \\ m: d_E^2 = 2m - \mu}}^{\min(d, m_c)} \Theta_{d_E, d, m, \mu} N^{m-\mu-d} \right]. \end{aligned} \quad (22)$$

The expression in (22) is just the expression in (20) with those terms that are correctable by EPC coding subtracted. By identifying the maximum exponent of the interleaver length N in (22) and (20), we can compare the asymptotic BER of stand-alone RSCC and EPCC-RSCC in the limit of large interleaver size. Assuming the minimum Hamming weight of the outer RSCC code is 2, we show the maximum interleaver gain exponent per d_E^2 for the nonprecoded dicode channel concatenated with RSCC and EPCC-RSCC ($d_c = 10, m_c = \{3, 4\}, L_c = 1$) in Table II for $d_E^2 = \{1, \dots, 4\}$, in Table III for $d_E^2 = \{5, 6\}$, and in Table IV for $d_E^2 = 7$. We also list for each d_E^2 , the generating error patterns and their corresponding parameters d, m , and μ . In addition, under each interleaver gain, we list in parenthesis the corresponding multiplicative coefficient $\Theta_{d_E, d, m, \mu}$, excluding the term $A^O(d)\bar{A}^I(d)$ relating to the outer RSCC Hamming error weight distribution.

For the precoded dicode channel concatenated to the RSCC, using the same approximations as above, it can be shown that the ML BER upper bound of Appendix A is dominated by the terms

$$P_b < \frac{1}{2K} \sum_{d_E=2}^{\infty} \sum_{d=2}^N \Theta_d N^{-\lceil \frac{d}{2} \rceil} e^{-\frac{d_E^2}{8\sigma^2}} \quad (23)$$

where Θ_d is given by

$$\Theta_d = A^O(d)\bar{A}^I(d) \left[\frac{d}{2} \right] \binom{d}{\lceil \frac{d}{2} \rceil}. \quad (24)$$

Note that we only kept those terms with error length L equal to the Hamming distance d^H , since they have the dominant interleaver gain exponent at each d^H . In this case, the dominant error will have $d_E^2 = d^H$ as shown in Table V.

TABLE IV
INTERLEAVER GAIN OF RSCC WITH NO PRECODING VERSUS EPCC-RSCC, $d_E^2 = 7$

$d_E^2 = 7$	Error pattern classes	RSCC	EPCC-RSCC
$m = 4$ $\mu = 1$ $d = 4 \rightarrow d_T$		N^{-1} (4)	$[N^{-1}]_{m_c=3}^{(4)}$ $[N^{-8}]_{m_c=4}^{(6237000)}$
$d_E^2 = 7$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-2} (6)	N^{-2} (6)
$d_E^2 = 7$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-2} (6)	N^{-2} (6)

TABLE V
INTERLEAVER GAIN OF RSCC WITH PRECODING VERSUS EPCC-RSCC, $d_E^2 = \{2, 3, 4, 5\}$

d_E^2	Error pattern classes	RSCC	EPCC-RSCC
$m = 1$ $\mu = 0$ $d^H = 2$ $L = 2$		N^{-1} (2)	$d^H = 11$ N^{-10} ($\frac{155925}{4}$)
$d_E^2 = 3$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 1$ $d^H = 3$ $L = 3$		N^{-2} (6)	$d^H = 11$ N^{-10} (779625)
$d_E^2 = 4$	Error pattern classes	RSCC	EPCC-RSCC
$m = 2$ $\mu = 0$ $d^H = 4$ $L = 4$		N^{-2} (12)	$d^H = 11$ N^{-9} ($\frac{779625}{2}$)
$d_E^2 = 5$	Error pattern classes	RSCC	EPCC-RSCC
$m = 3$ $\mu = 1$ $d^H = 5$ $L = 5$		N^{-3} (60)	$d^H = 2$ N^{-2} (1)

First, we note that for RSCC on the nonprecoded diode channel, the interleaver gain exponents are all negative for $d_E^2 = 1$ to $d_E^2 = 3$, which are the terms that dominate the BER for medium to high SNRs. Second, we note that the error patterns, for this same range of error Euclidean distances, up to $d = 10$, all belong to the dominant error class. As a result, EPCC-RSCC manages to substantially increase the interleaver gain by a factor of N^9 . Also, for $d_E^2 = 4$, where stand-alone RSCC does not achieve any interleaver gain, the EPCC-RSCC scheme has an impressive interleaver gain of N^{-9} .

The extremely low exponents suggest that EPCC-RSCC will have a large gain even for relatively short interleavers, and would thus deliver satisfactory gains in short to medium RSCC codeword sizes. At the same time, for such short interleavers, the stand-alone RSCC would considerably suffer in terms of the turbo gain. These conclusions will be numerically demonstrated in the next section by evaluating the BER bound for interleavers

as short as 100 bits. Furthermore, although $\Theta_{d_E, d, m, \mu}$ is significantly larger in EPCC-RSCC compared to stand-alone RSCC for the same d_E^2 , the term $\Theta_{d_E, d, m, \mu} N^{m-\mu-d}$ is still several orders of magnitude lower for EPCC-RSCC. Although less important, we also show the interleaver gain for higher error Euclidean distances in Tables III and Table IV. Most notably, EPCC-RSCC ($d_c = 10, m_c = 3, L_c = 1$) corrects errors belonging to the dominant error class for $d_E^2 = 5$ and $d_E^2 = 6$, increasing, in the process, the maximum interleaver gain by a factor of N , a turbo gain that becomes more substantial for large interleavers. Actually, for $d_E^2 = 6$, stand-alone RSCC possess no interleaver gain, while EPCC-RSCC's BER is dominated by the untargeted set of nondominant errors that result in the factor N^{-1} , still achieving an interleaver gain. On the other hand, EPCC-RSCC ($d_c = 10, m_c = 3, L_c = 1$) would offer no advantage when $d_E^2 = 7$. Note that although all errors belong to χ^{dom} when $m = 4$, their multiplicity m exceeds

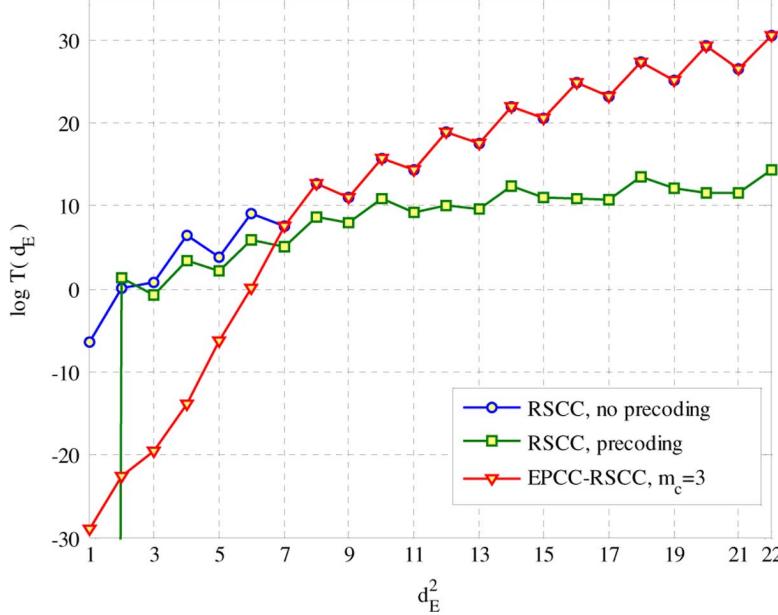


Fig. 5. $\log \bar{T}(d_E)$ of $(7, 5)$ RSCC with and without precoding versus $(7, 5)$ RSCC + $(630, 616)$ EPCC of multiple correction power $m_c = 3$.

the maximum multiple-error-pattern correction capability of $m_c = 3$. However, EPCC-RSCC ($d_c = 10, m_c = 4, L_c = 1$) manages to improve the maximum interleaver gain to N^{-2} , by reducing the contribution of χ^{dom} to $d_E^2 = 7$ by a factor of N^7 .

Comparing the interleaver gain exponents of the stand-alone RSCC with channel precoding and EPCC-RSCC in Table V, we note that EPCC-RSCC focuses on error events of $d_E^2 \leq 4$ and $d^H \leq 4$, improving the interleaver gain by a factor of N^9, N^8 , and N^7 at $d_E^2 = \{2, 3, 4\}$, respectively. Note that the RSCC's interleaver gain is better by a factor of N compared to EPCC-RSCC at $d_E^2 = 5$, suggesting that as SNR goes down, RSCC with precoding would have a better BER.

Overall, the comparison of the maximum interleaver gain factors predicts that the EPCC-RSCC's BER floor will be far lower than that of stand-alone RSCC with and without precoding. We summarize that EPCC-RSCC offers an effective way to enhance spectrum thinning of low Euclidean distances (and thus high SNR performance), whereas existing serial concatenation allows only a limited ability of lowering the error floor since the floor BER is not a strong function of the interleaver size in existing methods.

VI. NUMERICAL ANALYSIS AND SIMULATION RESULTS

Utilizing the analytic ML upper bound on the BER of stand-alone RSCC versus EPCC-RSCC, we study the relative performance of these systems for different levels of the severity of ISI. We make extensive use of (18) for this. We also study the dicode channel $1 - D$ for a variety of EPCC parameters. We will assume throughout the analysis that the SNR rate penalty (in dB) is proportional to $10 \log_{10} \frac{1}{R}$, where R is the code rate. Also recall that EPCC-RSCC needs to be employed without channel precoding.

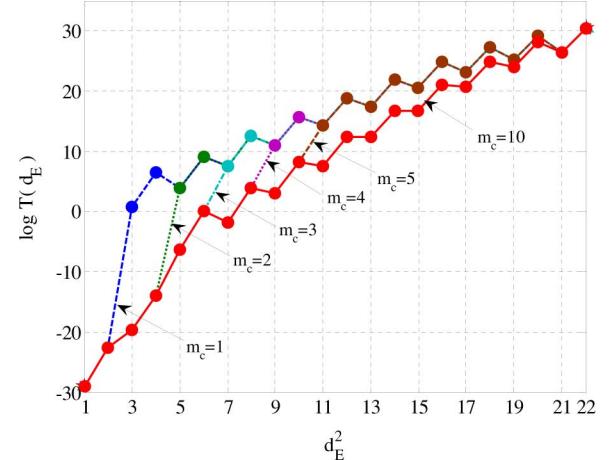


Fig. 6. $\log \bar{T}(d_E)$ of $(7, 5)$ RSCC + $(630, 616)$ EPCC for $m_c = \{1, 2, 3, 4, 10\}$.

A. The ML BER Bound and the Iterative Decoder Simulation BER

The \log of the average Euclidean distance distribution, $\log \bar{T}(d_E)$ calculated using (13), is shown in Fig. 5 for the nonprecoded dicode channel concatenated to either RSCC or EPCC-RSCC. Fig. 5 also includes the Euclidean distance distribution for the precoded dicode channel concatenated to RSCC, derived in a similar way to that of [14]. $\log \bar{T}(d_E)$ is calculated for an RSCC of code sequence length $N = 616$, which is also the interleaver size. The RSCC has generator polynomial connections $(7, 5)$ in octal format with base rate $1/2$ punctured to rate $R = 8/9$. The inner code is a $(630, 616)$ EPCC with $\{L_c = 1, m_c = 3, d_c = 10\}$.

From the average Euclidean distance distribution, we can conclude that for RSCC, precoding would exhibit larger interleaver gains compared to no precoding in the waterfall region,

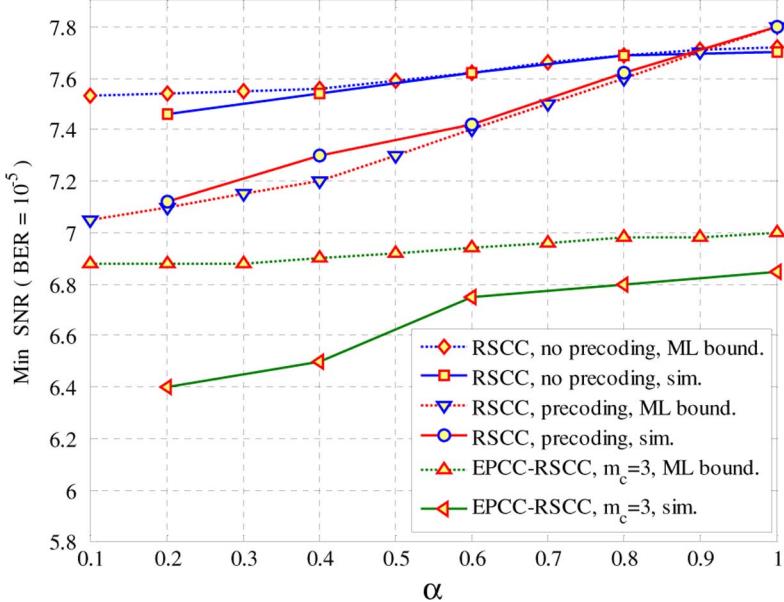


Fig. 7. SNR required to achieve a 10^{-5} BER: simulated (using TEs) versus bound for various concatenation systems using (7, 5) RSCC.

i.e., low to medium SNRs. This is because $\log \bar{T}(d_E)$ is lower for precoding everywhere when $d_E^2 \geq 3$. However, for higher SNRs, in the error floor region, the contribution of squared Euclidean distance 2 becomes stronger, and as seen in the figure, the average number of Hamming weight 2 errors that generate $d_E^2 = 2$ is more for the precoded case than it is for the nonprecoded case. On the other hand, the EPCC concentrates on low Euclidean distances, reducing their frequency substantially up to $d_E^2 = 6$. This would result in improved BER in the error floor and similar waterfall BER to the concatenation of nonprecoded dicode and stand-alone RSCC. Fig. 6 shows the truncated Euclidean distance distribution for different values of the EPCC correction power m_c , where we observe that if we continue to increase m_c beyond 3, the frequency suppression of Euclidean distances $d_E^2 \leq 6$ will remain the same. This means that the error floor will cease to improve after this point, and the gains will mostly be in the waterfall region. In Fig. 7 we study the upper bounds on the ML decoder BER along with the simulated BERs of the iterative decoders for $0 < \alpha \leq 1$. The code parameters are the same as in Fig. 5. Comparing the minimum SNR required for a 10^{-5} BER for the nonprecoded $1 - \alpha D$ channel, it is shown based on the bound that the ML decoder of EPCC-RSCC has a gain of about 0.7 dB compared to the ML decoder of stand-alone RSCC for $0 < \alpha \leq 1$. On the other hand, the ML decoder of RSCC with precoding has a gain of 0.5 dB at $\alpha = 0.1$ which turns into a loss of 0.1 dB at $\alpha = 1$, relative to RSCC without precoding. The ML BER of RSCC with precoding improves as $\alpha \rightarrow 0$, since the average Euclidean distance of dominant error patterns grows with their Hamming distance when $\alpha < 1$. This is because for $\alpha \neq 1$ crossing branches contribute to Euclidean distance growth (by $(1 - \alpha)^2 > 0$ for every crossing). On the other hand, for a given EPC power m_c , the ML BER of the EPCC-RSCC improves only slightly as $\alpha \rightarrow 0$, since the list of targeted dominant error patterns is the same for all α , and those are corrected irrespective of their Euclidean distances.

The simulated BERs of the respective iterative TEs are also shown in Fig. 7. To guarantee convergence of the suboptimal iterative decoders we run 10 turbo iterations for all TE systems. We used up to 100 test patterns in the list-decoder of the TE-EPCC [17], and each test word has up to two error patterns from the dominant list. One important remark regarding the comparison of the ML upper bound and simulation BER is that the bound is averaged over all possible interleaver permutations, i.e., uniform interleaving is assumed. This means that for certain interleaver instances, the ML BER can be above the upper bound. However, at least one interleaver will have the same BER as the average uniform interleaver [6], [13], and is guaranteed to fall below the upper bound. Another remark is that the upper bound is on the ML decoder BER, while the simulated BERs reflect the performance of iterative decoding and turbo equalization, which is suboptimal.

For these reasons, sometimes simulated performance will actually be worse than the bound, as can be seen in the case of RSCC with precoding. Nonetheless, the actual gain gaps between TE-EPCC and conventional TEs seem even larger than predicted by the ML upper bound. As such, we believe that the ML bound comparison provides a conservative estimate on the relative performance of our scheme. One final remark is that the EPCC-RSCC ML bound is somewhat loose at low levels of ISI. The reason is that the energies of dominant error events are much larger in long events compared to short ones, and this larger spread of energies means that the list-decoder's miscorrection probability is lower. This is because the list decoder can now order the candidate codewords more accurately according to their more widely spaced reliabilities.

In Fig. 8 we plot the SNR gain of EPCC-RSCC over stand-alone RSCC for the nonprecoded dicode channel defined as the difference in the minimum SNR required to achieve a BER of 10^{-7} for both systems. For this relatively low BER, there is a growing SNR gain furnished by EPCC-RSCC as $\alpha \rightarrow 1$, with the growth capped at 0.5 dB. Furthermore, for a given α , the

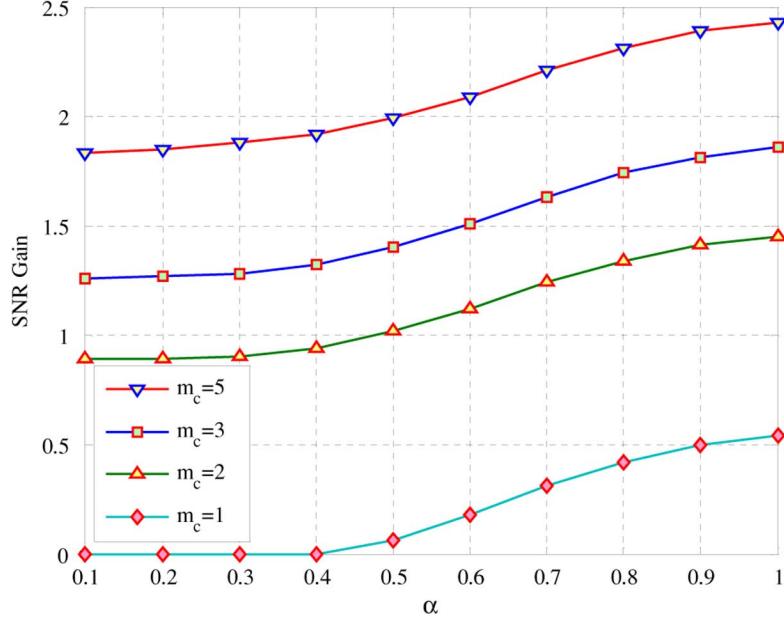


Fig. 8. SNR gain (dB) of EPCC-RSCC over RSCC with no precoding at a BER of 10^{-7} as a function of ISI level α for RSCC punctured rate $R = 8/9$ and interleave size $N = 616$.

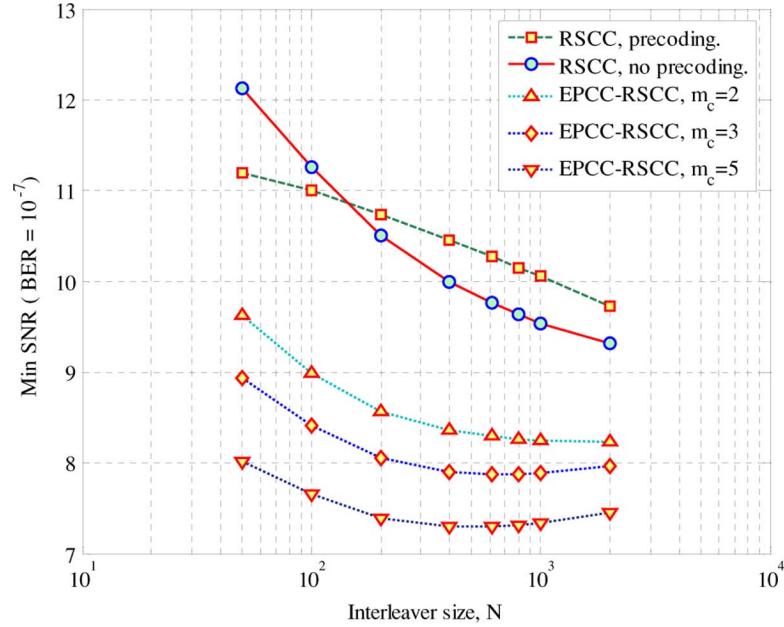


Fig. 9. Minimum SNR (dB) required to achieve a BER of 10^{-7} for RSCC with and without precoding versus EPCC-RSCC with $\{d_c = 10, L_c = 1\}$ for different interleave sizes. The punctured rate of RSCC is $8/9$ for all cases.

SNR gain of EPCC-RSCC grows with increasing m_c ; a 2 dB improvement can be achieved by increasing m_c from 1 to 5 for $0 < \alpha \leq 1$, which suggests that EPCC-RSCC can be made more resilient to severe ISI at the expense of higher decoder complexity.

B. Interleaver Gain

Based on the above discussion on the validity of using the ML upper bounds for performance comparison, we can investigate fairly low BERs, which are hard to reach through Monte Carlo simulation. In Fig. 9 we study the minimum SNR to achieve a low BER of 10^{-7} using EPCC-RSCC and stand-alone RSCC

with and without precoding and based on interleaver lengths $N = \{50, 100, 200, 400, 600, 800, 1000, 2000\}$ bits. Again, the outer RSCC is a $(7, 5)$ code punctured to rate $R = 8/9$.

The SNR gain of EPCC-RSCC over RSCC with no precoding using $N = 50$ is 2.3 dB, 3 dB, and 4 dB for EPC powers $m_c = \{2, 3, 5\}$, respectively. On the other hand, using $N = 2000$, this shrinks to 1.1 dB, 1.3 dB, and 1.8 dB for EPCC correction powers $m_c = \{2, 3, 5\}$, respectively.

We note that as the interleaver size N of EPCC-RSCC increases, the turbo gain of EPCC-RSCC increases accordingly. Also, since we maintain the same number of parity bits as the codeword length increases, less SNR rate penalty is incurred as

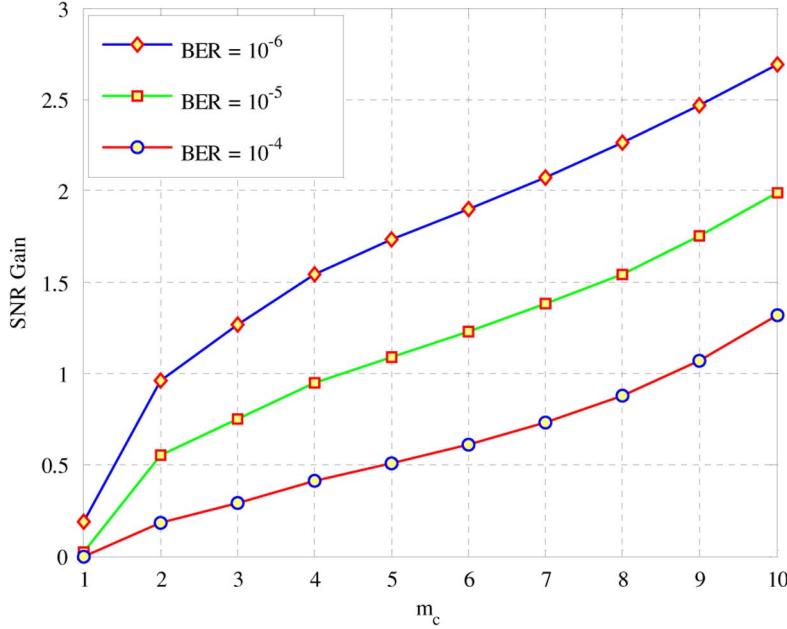


Fig. 10. SNR gain (dB) of EPCC-RSCC over stand-alone RSCC with no precoding at several BER operating points as a function of m_c for the outer (7, 5) RSCC code, punctured to $R = 8/9$ and interleave size $N = 616$.

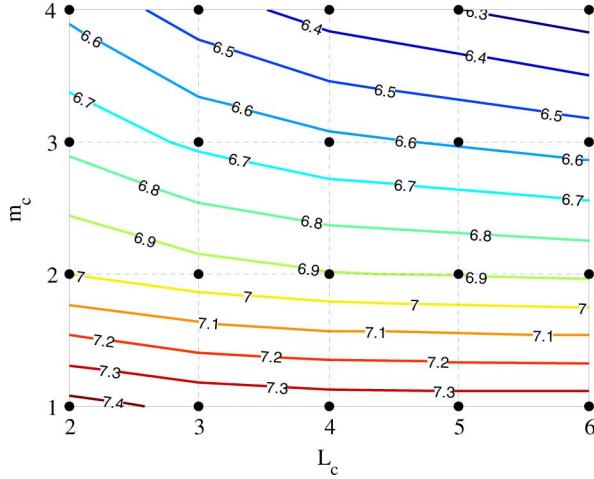


Fig. 11. An interpolated contour plot of the minimum SNR required by EPCC-RSCC to achieve $\text{BER} = 10^{-5}$ for different combinations of m_c and L_c , $N = 1200$, and RSCC (7, 5) of punctured $R = 8/9$.

N increases. On the other hand, the number of error patterns per codeword increases for larger N , surpassing EPCC's correction capability. Due to these conflicting effects of the EPCC-RSCC, its minimum required SNR plateaus and even increases as N increase slightly beyond a certain point. All in all, the relative advantage of EPCC-RSCC in practical systems would be most visible with small interleaver sizes. Furthermore, as can be observed in the figure, increasing m_c is also most effective for smaller interleaver sizes.

In practical EPCC construction, in order to obtain shorter EPCC code lengths while serially concatenating one EPCC codeword per RSCC interleaved codeword, i.e., $L_c = 1$, the EPCC code length is shortened from the (630, 616) EPCC at the same level of redundancy. When it is necessary to support interleaver sizes above 630, we duplicate EPCC codewords,

i.e., make $L_c > 1$, and use shortening to fit fractions of EPCC codewords in one interleaver block. For instance, we implement a (114, 100) EPCC of rate 0.88 for interleaver length $N = 100$, and a (630, 616) EPCC plus a shortened (398, 384) EPCC for $N = 1000$.

C. SNR Gain as Function of L_c and m_c

The performance of EPCC-RSCC can be further improved by increasing its multiple error correction capability m_c , per EPCC codeword. However, the complexity of the practical decoder would increase accordingly as more test words have to be constructed in the list decoder [17]. Fig. 10 shows EPCC-RSCC's SNR gains over stand-alone RSCC with no precoding for several BER operating points, as functions of $L_c = 1$ EPCC maximum correction capability (from $m_c = 1$ to $m_c = 10$). The RSCC is the same $R = 8/9$ code and $N = 1200$, and $d_c = 10$ for the EPCC. The curves demonstrate that EPCC-RSCC's SNR gain grows almost linearly as m_c is increased. Another design method to increase the correction capability of EPCC-RSCC, without considerably increasing its complexity, is to use $L_c > 1$ EPCC codewords per interleaver block. To study the design space spanned by m_c and L_c for a given interleaver size, we evaluate the BER bound for the set composed of the Cartesian product of the sets $m_c = \{1, 2, 3, 4\}$ and $L_c = \{2, 3, 4, 5, 6\}$. Then we plot a continuous contour of the minimum SNR to achieve $\text{BER} = 10^{-5}$ by interpolating the values found at the elements of the Cartesian product. A contour plot with an SNR step of 0.1 dB is shown in Fig. 11. Again, $N = 1200$ and $d_c = 10$. We note that the combinations $\{L_c = 2, m_c = 4\}$ and $\{L_c = 5, m_c = 3\}$ require a similar minimum SNR of 6.6 dB to achieve a BER of 10^{-5} . Nonetheless, as can be seen in Fig. 11, the slope of the equi-SNR contour lines decreases for higher L_c and lower m_c . This means that as the number of

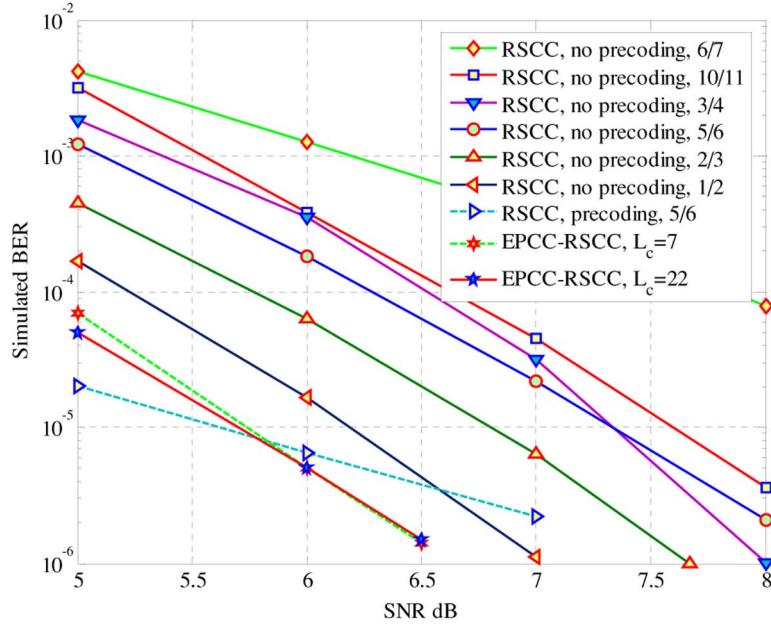


Fig. 12. BER simulation of EPCC-RSCC and stand-alone RSCC for various rates.

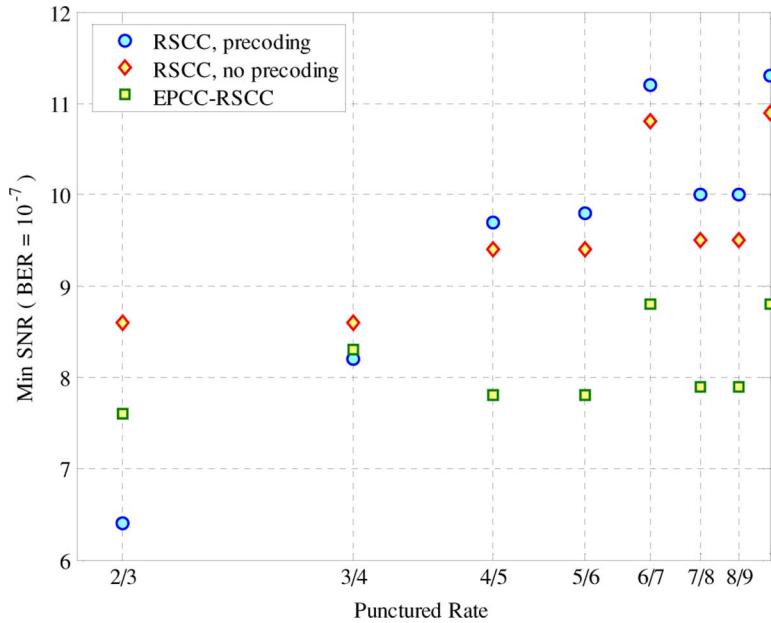


Fig. 13. Minimum SNR required to achieve 10^{-7} ML BER for stand-alone RSCC with and without precoding versus EPCC-RSCC, as a function of the outer (7, 5) RSCC punctured rate for $N = 1200$ and assuming a $\{m_c = 3, L_c = 1, d_c = 10\}$ EPCC.

EPCC codewords L_c increases per interleaver block, the correction capability plateaus, especially when $m_c \leq 2$. This is due to the higher level of redundancy required by the shortened EPCC to maintain the maximum correction capability m_c of the original nonshortened EPCC. For instance, at one extreme, to maintain the correction capability at a shortened EPCC code length of 44 bits, i.e., $L_c = 40$ and $N = 1200$, a shortened EPCC of rate 0.68 would incur a staggering rate penalty of 1.7 dB. An alternative concatenation approach that avoids the rate penalty of serial concatenation to a short inner EPCC is discussed in [25].

D. Puncturing Rate

We also wish to study the advantage of EPCC-RSCC at various total system rates and distributions of redundancy between the outer RSCC and inner EPCC codewords. In Fig. 12, we compare the simulated BERs of RSCC without precoding versus EPCC-RSCC for interleaver length $N = 4312$, different RSCC puncturing rates and EPCC parameters: $m_c = 3$ and $d_c = 10$. The results show that EPCC-RSCC composed of either $L_c = 7$ (630, 616) EPCC or $L_c = 22$ (210, 199) EPCC concatenated to a rate $\frac{5}{6}$ RSCC achieves the same BER in the error floor region.

Furthermore, either configuration outperforms comparable rate stand-alone RSCCs, with $L_c = 22$ EPCC-RSCC furnishing a gain of 1.5 dB with respect to the rate- $\frac{3}{4}$ RSCC at $\text{BER} = 10^{-6}$, and $L_c = 7$ EPCC-RSCC delivering a similar gain over the rate- $\frac{5}{6}$ RSCC. Moreover, either EPCC-RSCC scheme delivers a 1 dB SNR gain over the precoded RSCC of rate $\frac{5}{6}$.

For a complete investigation of a wide range of coding rates, we compare the minimum SNR required to achieve a BER of 10^{-7} (obtained from the bound) for punctured RSCC rates from $\frac{2}{3}$ to $\frac{9}{10}$, comparing EPCC-RSCC and the stand-alone RSCC with and without precoding. Such a comparison is shown in Fig. 13 for interleaver length $N = 1200$, different punctured-rates for RSCC, and assuming a $L_c = 1$ EPCC with $m_c = 3$ and $d_c = 10$. We conclude from the results that EPCC-RSCC delivers a uniform gain of 1.5 dB for puncturing rates above $\frac{2}{3}$. The abnormal peak in BER for puncturing rate $\frac{6}{7}$ is due to the particular choice of puncturing table. The reason why the rate- $\frac{2}{3}$ and rate- $\frac{3}{4}$ stand-alone RSCCs with precoding outperform EPCC-RSCC can be explained by examining $A^O(2)$ for these puncturing rates. It is shown in [24] using a similar approach to [18], that the outer RSCC does not generate Hamming weight 2 errors for these low puncturing rates, i.e., the minimum free distance of RSCC is $d_{\min}^f = 3$. Since the BER performance of stand-alone RSCC with precoding is dominated by these errors in the floor region, the corresponding BER is significantly improved, surpassing EPCC-RSCC at those rates. In summary, the stand-alone RSCC is more effective when the minimum Hamming distance of the outer code is larger than 2 for the precoded dicode channel. The stand-alone RSCC is less effective than EPCC-RSCC for punctured high rate codes.

VII. CONCLUSION

In this work, we have studied the BER of the serial concatenation of an EPCC and an interleaved RSCC over ISI channels as an alternative to a single RSCC with and without an inner precoder. To facilitate the study of system performance for a wide range of coding rates, interleaver sizes, and EPCC design parameters, we have derived an upper bound on the ML BER of EPCC-RSCC that is easy to evaluate and that scales well with system parameters. We have also shown how the EPCC enhances turbo equalization by reducing the frequency of error words of low Euclidean distance, which dominate the BER in the error floor region. Numerical results, calculated via the derived bound, indicate that EPCC-RSCC delivers substantial gain for short interleaver lengths compared to stand-alone RSCC for precoded and nonprecoded ISI channels, which suggests the TE-EPCC scheme is more attractive than the conventional TE for hardware implementation. Also, we have demonstrated that EPCC-RSCC provides a uniform gain of 1.5 dB over stand-alone RSCC for puncturing rates above 0.75, which indicates EPCC-RSCC is more suitable for high rate applications, such as magnetic and optical recording applications, while the stand-alone RSCC is a better choice for very low coding rates assuming precoding is employed.

APPENDIX A

A. Simple BER Bound Expressions for $L_c = 1$

1) *Stand-Alone RSCC*: As discussed above for the nonprecoded $1 - \alpha D$ channel, the Euclidean distance of a compound error event of multiplicity m and γ crossing branches is given by

$$d_E^2 = 4\alpha\gamma + (1 - \alpha)^2d + 2\alpha m - \mu\alpha^2.$$

where d is the Hamming distance of the compound error event. In order to compare the channel SNR of different levels of ISI α , we make use of the noise variance normalization

$$\tilde{\sigma}^2 = \frac{1 + \alpha^2}{2}\sigma^2$$

in which case the dicode channel has the base noise variance σ^2 . When the EPCC is turned off, the expression of the BER reduces to

$$\begin{aligned} P_b &\leq \frac{1}{K} \sum_{d_E=1}^{\infty} Q\left(\frac{d_E}{2\tilde{\sigma}}\right) \sum_{d=2}^{d_T} \frac{A^O(d)\bar{A}^I(d)}{\binom{N}{d}} \\ &\quad \times \sum_{\mu=0}^1 \sum_{\substack{m=1 \\ m: \gamma \geq 0, \gamma \in \mathbb{N}}}^d \binom{d-m}{\gamma} \\ &\quad \times \left(\frac{1}{2}\right)^{d-m} \binom{N-d}{m-\mu} \binom{d-1}{m-1} \\ \gamma &= \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1 - \alpha)^2 d}{4\alpha} \end{aligned} \quad (25)$$

A good approximation of the Q function that is accurate for a wide range of abscissa is borrowed from [26], and is given by

$$Q\left(\frac{d_E}{2\tilde{\sigma}}\right) \approx \frac{1}{12}e^{-\frac{d_E^2}{8\tilde{\sigma}^2}} + \frac{1}{4}e^{-\frac{d_E^2}{6\tilde{\sigma}^2}}.$$

Hence, P_b in (25) is composed of two terms as in

$$P_b(\tilde{\sigma}) = \frac{1}{12}\check{P}_b(2\sqrt{2}\tilde{\sigma}) + \frac{1}{4}\check{P}_b(\sqrt{6}\tilde{\sigma})$$

where

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{\infty} A^O(d)\bar{A}^I(d) \\ &\quad \times \sum_{\mu=0}^1 \sum_{m=1}^d \sum_{\gamma=0}^{d-m} \binom{d-1}{m-1} \binom{d-m}{\gamma} \left(\frac{1}{2}\right)^{d-m} \\ &\quad \times N^{m-\mu-d} \frac{d!}{(m-\mu)!} e^{-\frac{4\alpha\gamma+(1-\alpha)^2d+2\alpha m-\mu\alpha^2}{\tilde{\sigma}^2}}. \end{aligned} \quad (26)$$

Evaluating the summation over γ by utilizing the binomial identity we obtain

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \\ &\quad \times \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2d-\mu\alpha^2}{\tilde{\sigma}^2}} \left(1 + e^{-\frac{4\alpha}{\tilde{\sigma}^2}}\right)^d \\ &\quad \times \sum_{m=1}^d \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(\frac{2Ne^{-\frac{2\alpha}{\tilde{\sigma}^2}}}{1 + e^{-\frac{4\alpha}{\tilde{\sigma}^2}}}\right)^m \end{aligned} \quad (27)$$

After some algebraic manipulation, (27) simplifies to

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(N)^d} \\ &\times \sum_{\mu=0}^1 N^{1-\mu} e^{-\frac{1+\alpha^2(d-\mu)}{\tilde{\sigma}^2}} \frac{d!}{(1-\mu)!} \left(\cosh \left(\frac{2\alpha}{\tilde{\sigma}^2} \right) \right)^{d-1} \\ &\times {}_1F_1 \left(1-d; 2-\mu; -N \operatorname{sech} \left(\frac{2\alpha}{\tilde{\sigma}^2} \right) \right) \end{aligned} \quad (28)$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind [27], [28].

2) *EPCC-RSCC*: The BER of EPCC-RSCC, P_b^{epcc} , is expressed as the residual error rate after subtracting the error rate component that is correctable by the EPCC, P_b^C , from the stand-alone RSCC BER for the nonprecoded channel, P_b , as expressed in

$$P_b^{\text{epcc}}(\tilde{\sigma}) = P_b(\tilde{\sigma}) - P_b^C(\tilde{\sigma}).$$

Similar to (26), using the Q function approximation we have

$$P_b^C(\tilde{\sigma}) = \frac{1}{12} \check{P}_b^C(2\sqrt{2}\tilde{\sigma}) + \frac{1}{4} \check{P}_b^C(\sqrt{6}\tilde{\sigma})$$

with

$$\begin{aligned} \check{P}_b^C(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ &\times \sum_{m=1}^{\min(d, m_c)} \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right)^m \end{aligned} \quad (29)$$

which can be expanded into two sum terms depending on the value of d

$$\begin{aligned} \check{P}_b^C(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{m_c} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ &\times \sum_{m=1}^d \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right)^m \\ &+ \frac{1}{K} \sum_{d=m_c+1}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ &\times \sum_{m=1}^{m_c} \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right)^m. \end{aligned} \quad (30)$$

After some algebraic manipulation and gathering of geometric series terms we obtain the simplified expression

$$\begin{aligned} \check{P}_b^C(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ &\times 2N(d!) e^{-\frac{2\alpha}{\tilde{\sigma}^2}} {}_1F_1 \left(1-d; 2-\mu; -2Ne^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right) \\ &- \frac{1}{K} \sum_{d=m_c+1}^{\infty} \frac{A^O(d)\bar{A}^I(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} \\ &\times \frac{d! e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}}}{(m_c+1-\mu)!} \binom{d-1}{m_c} \left(2Ne^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right)^{m_c+1} \\ &\times {}_2F_2 \left(1, m_c+1-d; m_c+1, m_c+2-\mu; -2Ne^{-\frac{2\alpha}{\tilde{\sigma}^2}} \right) \end{aligned} \quad (31)$$

where ${}_2F_2$ is the generalized hypergeometric function [29].

3) *Stand-Alone RSCC and Precoded ISI Channels*: By examining the precoded trellis in Fig. 3, the squared Euclidean distance of a compound error event of Hamming distance d and length L for the precoded channel $\frac{1-\alpha D}{1 \oplus D}$ is

$$d_E^2 = \left\lceil \frac{d}{2} \right\rceil + \left\lfloor \frac{d}{2} \right\rfloor \alpha^2 + 4\alpha\gamma + (1-\alpha)^2(L-d).$$

Substituting this expression in the ML bound on BER of the precoded diode channel derived in [14], and utilizing the approximation of the Q function once again, we obtain

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^N \frac{A^O(d)\bar{A}^I(d)}{\binom{N}{d}} \\ &\times \sum_{L=d}^N \left(\frac{1}{2} \right)^{L-d} \binom{N-L+\lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor} \binom{L-1-\lceil \frac{d-1}{2} \rceil}{\lceil \frac{d-1}{2} \rceil} \\ &\times e^{-\frac{(1-\alpha)^2(L-d)+\lceil \frac{d}{2} \rceil+\lfloor \frac{d}{2} \rfloor}{\tilde{\sigma}^2} \alpha^2} \sum_{\gamma=0}^{L-d} \binom{L-d}{\gamma} e^{-\frac{4\alpha\gamma}{\tilde{\sigma}^2}}. \end{aligned} \quad (32)$$

Evaluating the summation over γ by utilizing the binomial identity, we obtain after some simplification

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^N \frac{A^O(d)\bar{A}^I(d)}{\binom{N}{d}} e^{-\frac{\lceil \frac{d}{2} \rceil+\lfloor \frac{d}{2} \rfloor}{\tilde{\sigma}^2} \alpha^2} \\ &\times \sum_{L=d}^N \binom{N-L+\lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor} \\ &\times \binom{L-1-\lceil \frac{d-1}{2} \rceil}{\lceil \frac{d-1}{2} \rceil} \Psi(\alpha, \tilde{\sigma})^{L-d} \\ \Psi(\alpha, \tilde{\sigma}) &= e^{-\frac{1+\alpha^2}{\tilde{\sigma}^2}} \cosh \left(\frac{2\alpha}{\tilde{\sigma}^2} \right) \end{aligned} \quad (33)$$

This can be simplified to a single sum over d by the utility of the generalized hypergeometric representation, which is given by

$$\check{P}_b(\tilde{\sigma}) \leq \frac{1}{K} \sum_{d=2}^N \frac{A^O(d)\bar{A}^I(d)}{\binom{N}{d}} \binom{N - \lceil \frac{d}{2} \rceil}{\lfloor \frac{d}{2} \rfloor} e^{-\frac{\lceil \frac{d}{2} \rceil + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \\ \times {}_3F_2 \left(1, \left\lfloor \frac{d+1}{2} \right\rfloor, d-N; \left\lceil \frac{d}{2} \right\rceil - N, 1; \Psi(\alpha, \tilde{\sigma}) \right). \quad (34)$$

When $N \gg d$, we can use similar approximations to the ones used in the derivation of the interleaver gain exponent, by which one reaches a looser, albeit simpler, bound

$$\check{P}_b(\tilde{\sigma}) \leq \frac{1}{K} \sum_{d=2}^N A^O(d)\bar{A}^I(d) \frac{d!}{\lceil \frac{d}{2} \rceil!} N^{-\lceil \frac{d}{2} \rceil} e^{-\frac{\lceil \frac{d}{2} \rceil + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \\ \times {}_3F_2 \left(1, \left\lfloor \frac{d+1}{2} \right\rfloor, d-N; \left\lceil \frac{d}{2} \right\rceil - N, 1; \Psi(\alpha, \tilde{\sigma}) \right). \quad (35)$$

REFERENCES

- [1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE ICC*, Geneva, Switzerland, May 1993, pp. 1740–1745.
- [2] C. Douillard, M. Juel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507–511, Sep./Oct. 1995.
- [3] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 67–80, Jan. 2004.
- [4] M. Tuchler, R. Koetter, and A. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.
- [5] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.
- [6] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 409–428, Mar. 1996.
- [7] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.
- [8] W. Ryan, "Performance of high rate turbo codes on a PR4-equalized magnetic recording channel," in *Proc. IEEE ICC*, Jun. 1998, vol. 2, pp. 947–951.
- [9] M. Reed, C. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1693–1699, Dec. 1998.
- [10] T. V. Souvignier, M. Oberg, P. H. Siegel, R. E. Swanson, and J. K. Wolf, "Turbo decoding for partial response channels," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1297–1308, Aug. 2000.
- [11] L. McPheters, S. McLaughlin, and K. Narayanan, "Precoded PRML, serial concatenation, and iterative (turbo) decoding for digital magnetic recording," *IEEE Trans. Magn.*, vol. 35, no. 5, pp. 2325–2327, Sep. 1999.
- [12] J. Li, K. Narayanan, E. Kurtas, and C. Georghiades, "On the performance of high-rate TPC/SPC codes and LDPC codes over partial response channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 723–734, May 2002.
- [13] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.
- [14] M. Oberg and P. H. Siegel, "Performance analysis of turbo-equalized partial response channels," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 436–444, Mar. 2001.
- [15] J. Moon and J. Park, "Detection of prescribed error events: Application to perpendicular recording," in *Proc. IEEE ICC*, May 2005, vol. 3, pp. 2057–2062.
- [16] J. Park and J. Moon, "High-rate error-correction codes targeting dominant error patterns," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2573–2575, Oct. 2006.
- [17] H. Alhussien, J. Park, and J. Moon, "Iterative decoding based on error pattern correction," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 181–186, Jan. 2008.
- [18] D. N. Rowitch, "Convolutional and turbo coded multicarrier direct sequence CDMA, and applications of turbo codes to hybrid ARQ communication systems," Ph.D. dissertation, Univ. California San Diego, La Jolla, CA, Jun. 1998.
- [19] J. Park and J. Moon, "A new class of error-pattern-correcting codes capable of handling multiple error occurrences," *IEEE Trans. Magn.*, vol. 43, no. 6, pp. 2268–2270, Jun. 2007.
- [20] J. Park and J. Moon, "Error-pattern-correcting cyclic codes tailored to a prescribed set of error cluster patterns," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1747–1765, Apr. 2009.
- [21] J. Hagenauer and P. Hoeher, "A viterbi algorithm with soft-decision outputs and its applications," in *Proc. IEEE GLOBECOM*, Nov. 1989, vol. 3, pp. 1680–1686.
- [22] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.
- [23] M. Oberg and P. H. Siegel, "Performance bound for parity-check coded partial-response channels," in *Proc. IEEE ICC*, 2001, vol. 9, pp. 2701–2705.
- [24] H. Alhussien, "Channel matched iterative decoding for magnetic recording systems," Ph.D. Dissertation, Univ. Minnesota, Twin Cities, Minneapolis, MN, Apr. 2009.
- [25] H. Alhussien and J. Moon, "An iteratively decodable tensor product code with application to data storage," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 2, pp. 228–240, Feb. 2010.
- [26] M. Chiani, D. Dardari, and M. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 840–845, Jul. 2003.
- [27] M. Petkovsek, H. S. Wilf, and D. Zeilberger, *A = B*. London, U.K.: A K Peters, Ltd., 1997.
- [28] L. J. Slater, *Confluent Hypergeometric Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1960.
- [29] L. J. Slater, *Generalized Hypergeometric Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1966.

Hakim Alhussien (S'05–M'09) received the B.S. and M.S. degrees in electrical engineering with highest honors from Jordan University of Science and Technology (JUST), Irbid, in 2001 and 2003, respectively, and the M.S.E.E. and Ph.D. degrees in electrical engineering from the University of Minnesota Twin-Cities, Minneapolis, in 2008 and 2009, respectively.

From 2003 to 2004, he was an Instructor with the Department of Electrical Engineering, University of Yarmouk, Jordan. From 2004 to 2008, he held Research and Teaching Assistant positions with the Department of Electrical and Computer Engineering, University of Minnesota. Since September 2008, he has been a Staff Systems Engineer with Link-A-Media Devices, Santa Clara, CA. His main research interests are in the applications of coding, signal processing, and information theory to data storage systems.

Jaekyun Moon (S'89–M'90–SM'97–F'05) received the B.S.E.E. degree with high honors from Stony Brook University, Stony Brook, NY, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA.

From 1990 through early 2009, he was with the faculty of the Department of Electrical and Computer Engineering at the University of Minnesota, Twin Cities. He was also Consulting Chief Scientist at DSPG, Inc., from 2004 to 2007, and Chief Technology Officer at Link-A-Media Devices Corp. in 2008. He is now a Professor of Electrical Engineering at Korea Advanced Institute of Science and Technology (KAIST), Daejeon. His research interests are in the area of channel characterization, signal processing, and coding for data storage and digital communication.

Prof. Moon received the 1994–1996 McKnight Land-Grant Professorship from the University of Minnesota. He also received the IBM Faculty Development Awards as well as the IBM Partnership Awards. He was awarded the National Storage Industry Consortium (NSIC) Technical Achievement Award. He served as Program Chair for the 1997 IEEE Magnetic Recording Conference. He is also Past Chair of the Signal Processing for Storage Technical Committee of the IEEE Communications Society. In 2001, he cofounded Bermai, Inc., a fabless semiconductor start-up, and served as founding President and CTO. He served as a Guest Editor for the 2001 IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS issue on Signal Processing for High Density Recording. He also served as an Editor for IEEE TRANSACTIONS ON MAGNETICS in the area of signal processing and coding for 2001–2006.