Serial Concatenated TCM With an Inner Accumulate Code—Part II: Density-Evolution Analysis

Hugo M. Tullberg, *Member, IEEE,* and Paul H. Siegel, *Fellow, IEEE*

***Abstract*—In a companion paper, we showed the existence of decoding thresholds for maximum-likelihood (ML) decoding of a serial concatenated trellis-coded modulation (SCTCM) system with one or more inner accumulate codes. In this paper, we compute the decoding thresholds for an iterative, non-ML decoder by density evolution (DE), assuming infinite blocklengths. We also derive a stability condition for the particular case of an outer parity-check code and a single inner accumulate code. We show that, for equiprobable signaling, the bit-wise log-likelihood ratio densities for higher order constellations are symmetric. Furthermore, when used in DE, these densities can be averaged without significantly affecting the resulting threshold values. For an outer single parity-check code, the lowest decoding thresholds are achieved with two inner accumulate codes. For an outer repeat code, a single inner accumulate code gives the lowest thresholds. At code rates $r_c > 2/3$, the decoding thresholds for the SCTCM system are within 1 dB of the constellation-constrained channel capacity for additive white Gaussian noise channels, and within 1.5 dB for independent, identically distributed Rayleigh channels. Simulation results verify the computed thresholds.**

***Index Terms*—Accumulate codes, density evolution (DE), fading channels, iterative decoding, serial concatenation, stability condition, trellis-coded modulation (TCM).**

I. Introduction

TURBO codes [1] and interleaved serially concatenated codes [2] allow us to approach Shannon's theoretical capacity limit using practical, suboptimal iterative decoding. Trellis-coded modulation (TCM) [3] is a well-established technique to obtain coding gains without bandwidth expansion. The desire to design systems that provide bandwidth-efficient communication close to capacity motivated researchers to merge turbo-like codes and TCM, and several such "Turbo-TCM" systems have been proposed in the literature. Parallel concatenated TCM (PCTCM) was studied in [4] and [5], and serial concatenated TCM (SCTCM) was introduced in [6]. In order to reduce the decoding complexity, SCTCM schemes with low-complexity rate-1 inner codes were proposed in [7].

Inspired by the analytical tractability of repeat-accumulate (RA) codes [8] and their generalizations [9], this paper studies an SCTCM system with inner rate-1 accumulate codes. In this paper, we propose an SCTCM scheme with a single or multiple inner rate-1 accumulate code(s), each preceded by an interleaver, followed by a mapping to a higher order, Gray-labeled signal constellation.

To achieve higher spectral efficiency, we consider parity-accumulate (PA) codes, where the outer code is a single parity-check (SPC) code. However, the minimum distance of the outer SPC is $d_{\min}^{(o)} = 2$, and two or more inner accumulate codes are required to make the word-error rate (WER) go to zero as the blocklength tends to infinity. With an outer high-rate convolutional code with $d_{\text{free}}^{(o)} \geqslant 3$, we simultaneously achieve high spectral efficiency and asymptotically vanishing WER with only a single inner accumulate code.

Bit-interleaved coded modulation (BICM) was initially proposed to increase the diversity over fading channels, with only a modest performance degradation over additive white Gaussian noise (AWGN) channels [10]. BICM with iterative decoding (BICM-ID) has been shown to give almost the same performance as Turbo-TCM over AWGN channels, but at a lower complexity [11], [12]. A BICM system can be considered to be a special case of the proposed SCTCM system, since a BICM system consists of an outer convolutional code, no inner code, a channel interleaver, and a Gray-labeled constellation. We show a comparison of the performance of the proposed SCTCM system and BICM-ID over a correlated Rayleigh fading channel in [13] and [14].

In this paper, we restrict the binary outer code to be either an SPC code or a repeat code for the sake of simplicity of analysis. We present an asymptotic analysis of this SCTCM scheme under iterative decoding, where the blocklength goes to infinity, based on the density-evolution (DE) technique [15]. In a companion paper [14], we studied the performance of the same SCTCM scheme via union bounds (assuming the optimal and prohibitively complex maximum-likelihood (ML) decoding), and calculated the potential thresholds under ML decoding. However, due to the weakness of the union-bound technique, the thresholds calculated in this paper are much more encouraging, and indicate the potential of these coding schemes, even under suboptimal iterative decoding.

DE was introduced as a method to analyze message-passing decoders [15]. A code can be described by a graph, and the message-passing decoder passes messages along the edges of the graph. In DE, the messages are the densities of the log-likelihood ratios (LLRs) for the bits, and by tracing the evolution of

Paper approved by W. C. Kwong, the Editor for Optical Networks of the IEEE Communications Society. Manuscript received August 6, 2002; revised March 20, 2003. This work was supported in part by the Swedish Defence Research Agency (FOI), in part by the National Science Foundation under Grant NCR-9612802, in part by the Center for Wireless Communications at the University of California, San Diego, in part by Ericsson Corporation, and in part by the UC Discovery Grant Program. This paper was presented in part at the IEEE Vehicular Technology Conference, Atlantic City, NJ, October 2001.

H. M. Tullberg was with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407 USA. He is now with the Department of Communication Systems, Swedish Defence Research Agency (FOI), SE-581 11 Linköping, Sweden (e-mail: hugo.tullberg@foi.se).

P. H. Siegel is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407 USA (e-mail: psiegel@ucsd.edu).

Digital Object Identifier 10.1109/TCOMM.2004.841986

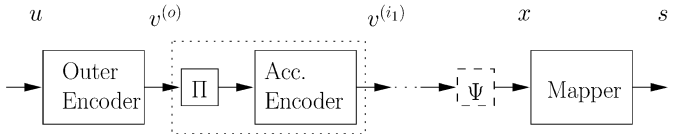

Fig. 1. SCTCM system under consideration consists of an outer repeat or SPC code, one or more inner interleaved accumulate codes, an optional channel interleaver Ψ, and a mapper to a Gray-labeled constellation.

the densities as the decoding progresses, we can determine the minimum channel signal-to-noise ratio (SNR) needed for the message-passing decoder to converge to the correct codeword with high probability. By applying DE, we find the threshold γ^* for a message-passing decoder for the SCTCM system under consideration.

For PA[1] codes, P_b decreases very slowly as the number of decoding iterations increases. It is, therefore, hard to determine an accurate numerical value for the threshold for PA[1] codes via DE. In this case, we derive a stability condition, similar to the one in [16], to determine thresholds for these codes. In the stability analysis, we assume that decoding has progressed successfully to a point where only a small fraction of bits remain in error. The stability condition then gives the minimum channel SNR required to ensure that the remaining errors vanish as the number of decoding iterations goes to infinity.

The probability density functions (pdfs) for the LLRs of the bits play an important role in the DE and stability analysis. Multiple bits are transmitted in each channel symbol, and the pdfs for the bits are different. We show that the resulting pdfs are symmetric [16] under certain assumptions. Furthermore, our analysis shows that the pdfs can be averaged, simplifying the analysis, without significantly affecting the computed thresholds.

We report numerical values for the thresholds for RA^m and PA^m codes of various rates, used with 8-ary phase-shift keying (8-PSK) and 16-quadrature amplitude modulation (QAM) constellations over AWGN and independent, identically distributed (i.i.d.) Rayleigh fading channels. The simulated performance of the proposed SCTCM system corresponds well to the computed thresholds.

The paper begins with a brief description of the proposed system in Section II. Section III treats the DE, and in Section IV, we derive the stability condition. In Section V, we report some simulation results. Our conclusions are presented in Section VI.

II. System Description

The encoder, shown in Fig. 1, consists of an outer block code, one or more inner interleaved accumulate codes, an optional channel interleaver, and a mapping to a higher order Gray-labeled signal constellation.

The outer code is either a $r_c = 1/n$ repeat code or a $r_c = (n-1)/n$ SPC code. The component codes are separated by an interleaver Π. In the analysis, we assume that Π is a uniform interleaver [17], and in the simulations, we use an S-random interleaver [18]. The accumulate code can be thought of as a recursive rate-1 convolutional code with generator matrix

$$G(D) = \left(\frac{1}{1 \oplus D}\right).$$

The memoryless mapper maps an m-tuple of bits, x, to a constellation point $s \in \mathcal{S}$, where $\mathcal{S}$ is a Gray-labeled constellation of size $|\mathcal{S}| = M = 2^m$.

We receive $r = \rho s + z$, where ρ is a sample from an i.i.d. Rayleigh fading process, and z is AWGN with variance σ^2 in each dimension. At the decoder, bit metrics are calculated by

$$P[x^i = b; O] = \sum_{s \in \mathcal{S}_b^i} P[r|s] \tag{1}$$

where x^i is the ith bit in the binary label of the transmitted symbol, and $\mathcal{S}_b^i = \{s \in \mathcal{S} | \ell^i(s) = b\}$ is the set of points in the constellation $\mathcal{S}$ such that the ith bit, $i \in \{1, \ldots, m\}$, in the binary label of the point s has the value b, $b \in \{0, 1\}$. The bit metrics are fed to a message-passing decoder, which is described in detail in Section III-C.

III. Density Evolution

In this section, we compute the LLR densities of the received bits for higher order constellations, and show that the densities are symmetric in the sense of (5) below [16]. (This property was earlier called "consistency" [19].) We describe the code graphs and the message-passing algorithm, and use DE to determine thresholds for the proposed SCTCM system when a message-passing decoder is used.

A. Densities for Higher Order Constellations

For output symmetric channels [15], it can be assumed that the all-zeros codeword was transmitted. All received bits will then have a common LLR density, which simplifies the DE analysis. Since we are considering higher order constellations, we cannot make this assumption. Instead, we assume that all constellation symbols are transmitted equally likely, and define the LLR as the LLR of receiving the same bit value as was transmitted. There are m bits in the label for each constellation symbol, so we get m possibly different LLR pdfs. The assumption that all constellation symbols are transmitted equally likely implies that the bits in the label will take on the values 0 and 1 with equal probability, and for long codewords, the LLR pdfs are independent of the transmitted codeword.[1]

Let $\mathcal{S}$ be a signal constellation of size $|\mathcal{S}| = M = 2^m$ admitting a binary labeling with m bits. Consider the ith bit of the binary label. (Henceforth, we omit the bit index i for clarity, but all functions depend on the bit under consideration.) Let $\mathcal{S}_0$ be the subset of $\mathcal{S}$ where the ith bit in the label is zero, and $\mathcal{S}_1 = \mathcal{S} \setminus \mathcal{S}_0$. Let η_0 (η_1) be the index set for the elements in $\mathcal{S}_0$ ($\mathcal{S}_1$), and let $s \in \mathcal{S}$ be the transmitted symbol.

We compute the LLR of receiving the same bit value as was transmitted. Since x depends on which symbol s was transmitted, we express LLR l as

$$l = g(r) = \begin{cases} g_0(r) = \log \dfrac{\sum_{j \in \eta_0} f_R(r|s_j)}{\sum_{j \in \eta_1} f_R(r|s_j)}, & s \in \mathcal{S}_0 \\ g_1(r) = \log \dfrac{\sum_{j \in \eta_1} f_R(r|s_j)}{\sum_{j \in \eta_0} f_R(r|s_j)}, & s \in \mathcal{S}_1 \end{cases} \tag{2}$$

where the random variable R is the received value, and $f_R(r|s_j)$ is the pdf of R, given that the symbol s_j was transmitted. If

[1]This assumption is similar to the channel adapter in [20].

we assume that all symbols are transmitted equally likely, the density for R becomes

$$f_R(r) = \frac{1}{M}\sum_{j=1}^{M} f_R(r|s_j). \tag{3}$$

Note that for any labeling $g_0(r) = -g_1(r)$.

The value of the LLR, $l = g(r)$, is a one-dimensional (1-D) random variable and a function of the possibly multidimensional random variable R. For 2-D modulation, let $r = (r_1, r_2)$. The LLR L then becomes a function of a 2-D random variable, $l = g(r_1, r_2)$. The pdf for the LLR, $f_L(l)$, can then be calculated by [21, p. 167]

$$f_L(l) = \int_{-\infty}^{\infty} \left.\frac{f_{R_1R_2}(r_1, r_2)}{\left|\frac{\partial g(r_1,r_2)}{\partial r_1}\right|}\right|_{r_1=g^{-1}(l,r_2)} dr_2. \tag{4}$$

If, for a fixed value of l and r_2, the equation $r_1 = g^{-1}(l, r_2)$ (or equivalently, $l = g(r_1, r_2)$) has multiple solutions, then (4) should be integrated over all such solutions.

In the 1-D case, the LLR functions g_0 and g_1 are functions of one random variable, and the partial derivative in (4) becomes the ordinary derivative of $g_0(r)$ or $g_1(r)$ with respect to r.

For binary phase-shift keying (BPSK) modulation, $s_1 = -s_0$ and $g_0 = 2s_0 r/\sigma^2$. It is straightforward to show that $f_L(l)$ is Gaussian with mean $\mathrm{E}[L] = 2s_0^2/\sigma^2$ and variance $\mathrm{var}[L] = 4s_0^2/\sigma^2 = 2\mathrm{E}[L]$. For BPSK modulation over a Rayleigh fading channel, the pdf conditioned on the fading power is Gaussian, and the unconditional pdf is given by

$$f_L(l|s) = \frac{e^{\frac{1}{2}\left(l-\sqrt{2+\frac{s^2}{\sigma^2}}\sqrt{\frac{l^2\sigma^2}{s^2}}\right)}\sigma}{2s\sqrt{2+\frac{s^2}{\sigma^2}}}.$$

For higher order constellations, we find the density in (4) numerically. For fading channels, we use (4) as the expression for the pdf conditioned on the fading power. The unconditional pdf is found by numerically integrating over the fading power distribution.

A density is called *symmetric* if [16]

$$f_L(-l) = e^{-l} f_L(l). \tag{5}$$

This symmetry property is used in the stability condition in Section IV.

Theorem 1: For a higher order constellation $\mathcal{S}$ and equiprobable signaling, the bit log-likelihood density $f_L(l)$ is symmetric in the sense of (5) for all bits in the labels.

Proof: The proof is given in the Appendix. ∎

B. Code Graphs

A linear code can be thought of as a linear system of equations, with variables and check equations. A graph can be used to visualize the structure of the code, i.e., the connections between variables (codeword bits and, if the code has memory, states) and parity-check equations.

A graph $G(V, E)$ is an ordered pair of disjoint sets V and E, where V is the set of vertices (or nodes) and E is the set of edges. If G is a graph, $V(G)$ denotes the set of vertices, and $E(G)$ the set of edges of G. A graph is said to be r-partite with

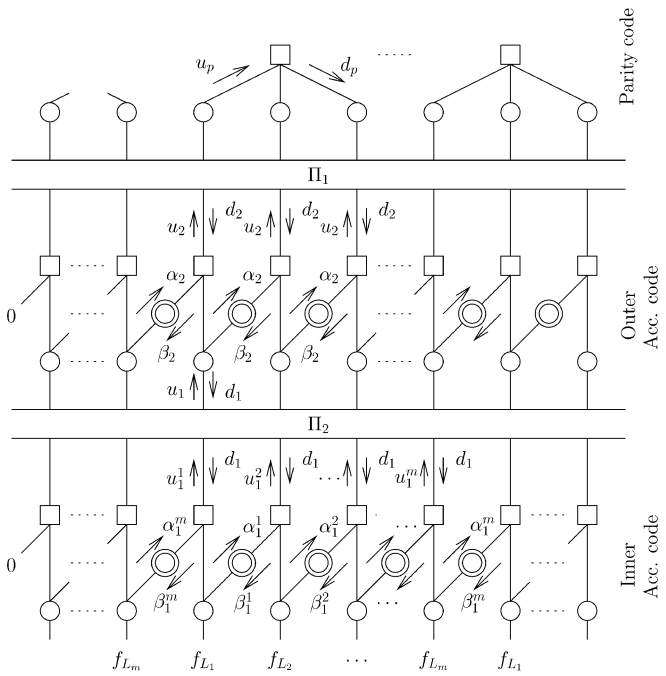

Fig. 2. Code graph for a $r_c = 2/3$ PA^2 code without channel interleaver Ψ.

r vertex classes $V_1, V_2, \ldots, V_r$ if the set of vertices can be partitioned $V(G) = V_1 \cup V_2 \cup \ldots \cup V_r$ and $V_i \cap V_j = \emptyset, \forall i \neq j$, and no edge joins two vertices from the same class. For code graphs, V consists of variable nodes corresponding to the codeword bits, check nodes corresponding to the parity-check equations and, if the code has memory, state nodes. Block codes, such as low-density parity-check (LDPC) codes [22], are described by bipartite (2-partite) graphs. We view the accumulate code as a convolutional code, so the graph representation of the accumulate code will have state nodes. Therefore, the graph representation of the RA^m and PA^m codes are 3-partite.

Two vertices $v, w \in V(G)$ are said to be adjacent if they are connected with an edge $e_{vw} \in E(G)$. The set of vertices adjacent to v is called the neighborhood (of depth 1) of v and denoted $N(v)$. The neighborhood of depth t of v is the set of vertices connected to v by a path of length at most t, and is denoted $N^t(v)$. Let E_v denote the set of edges connecting the vertex v to its neighborhood (of depth 1). The degree of a vertex v is the number of edges connected to that vertex, $\deg v = |E_v|$.

In Fig. 2, we show a code graph for a PA^2 code. Bit nodes are drawn as circles, state nodes as double circles, and parity-check nodes as squares.[2]

C. Message-Passing Decoders

In a message-passing decoder, messages are passed along the edges of the code graph. The message passed along an edge is the current opinion, or belief, about a bit. The message sent from a node v along an edge e, $\mathsf{m}(e)$, depends on all messages to node v except the message along the edge e. The code graph G

[2]An accumulate code has two states, 0 and 1, and the ending state in a trellis transition is the same as the output from the code. A separate state node, as in Fig. 2, is, therefore, somewhat superfluous, but highlights the memory of the code.

per se is undirected, but the message going from v to w, where $v, w \in V(G)$, is different from the message going from w to v.

Bit nodes represent bits, and all messages should agree on a common bit value. The outgoing message along the jth edge from a bit node of degree n is given by the sum of the incoming LLRs, except the contribution on the jth edge

$$\mathsf{m}_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \mathsf{m}_i. \tag{6}$$

A repeat code imposes the same constraint as a bit node—all variables should agree on a common bit value. The operation carried out in the "repeat node" is, therefore, the sum in (6). The final decision in the repeat node is based on all n incoming messages.

At the parity-check nodes, random variables are added under a parity constraint. The outgoing message on the jth edge from a check node of degree n is given by the tanh operation [23]

$$\tanh\left(\frac{\mathsf{m}_j}{2}\right) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \tanh\left(\frac{\mathsf{m}_i}{2}\right).$$

In the context of DE, the messages are not the beliefs about a particular received bit, but rather the pdf of the LLR of the bit. When we want to stress that the message on an edge e, $\mathsf{m}(e)$, is a pdf, we write $f_e(l)$.

At the bit and repeat nodes, random variables are added, and their pdfs are convolved. The outgoing message along the jth edge from a bit or repeat node of degree n is given by the $(n-1)$-dimensional convolution

$$f_{e_j}(l) = \bigotimes_{\substack{i=1 \\ i \neq j}}^{n} f_{e_i}(t_i).$$

At a parity-check node of degree n, the outgoing message on the jth edge is given by the $(n-1)$-dimensional integral

$$f_{e_j}(l) = \int_A \prod_{\substack{i=1 \\ i \neq j}}^{n} f_{e_i}(t_i) d\boldsymbol{t} \tag{7}$$

where the region A is given by

$$A = \left\{ \boldsymbol{t} \in \mathbb{R}^{n-1} : \tanh\left(\frac{l}{2}\right) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \tanh\left(\frac{t_i}{2}\right) \right\}.$$

We denote the operation in (7) by $\boxplus$.

In Fig. 2, we indicate the messages passed in the decoding of a PA^2 code. If no channel interleaver Ψ is present between the inner code and the signal mapper, then m different channel densities f_{L_i}, $1 \leqslant i \leqslant m$, are passed to the first accumulate code decoder. The forward and backward state messages are denoted α and β, and u and d denote the messages up to and down from the next decoder. We get m different α, β, and u messages in the first decoder (subscripts denote the code layer, and superscripts denote variables within the code). Due to the interleaver Π, the m upward messages u_1^i are averaged before passing to the next accumulate code decoder. From Fig. 2, we get the update equations for the inner accumulate code as

$$\begin{aligned} \alpha_1^i &= \left(\alpha_1^{i-1} \boxplus d_1\right) \otimes f_{L_i} \\ \beta_1^i &= \left(\beta_1^{i+1} \otimes f_{L_{i+1}}\right) \boxplus d_1 \\ u_1^i &= \left(\beta_1^i \otimes f_{L_i}\right) \boxplus \alpha_1^{i-1} \\ u_1 &= \frac{1}{m} \sum_{i=1}^{m} u_1^i \\ d_1 &= (\alpha_2 \boxplus d_2) \otimes \beta_2 \end{aligned} \tag{8}$$

where the index additions and subtractions are modulo m. Note that these are local updates, we do not perform the full forward and backward recursions of the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [24].

If multiple accumulate codes are used, the update equations for the jth accumulate code are the same as above, but the m channel densities f_{L_i} are replaced by the single message u_{j-1}, and the superscript i in (8) is omitted. The update equations for a $r_c = (n-1)/n$ SPC code and a $r_c = 1/n$ repeat code are

$$\begin{aligned} d_p &= u_p^{\boxplus(n-1)} \\ d_r &= u_r^{\otimes(n-1)} \end{aligned}$$

respectively, where the superscript $\boxplus(n-1)$ denotes applying the operation $\boxplus$ to $n-1$ arguments, and similarly for $\otimes(n-1)$.

If the interleaver Ψ is present, the channel densities are averaged over the m bits in the constellation, and we pass the density

$$f_L(l) = \frac{1}{m} \sum_{i=1}^{m} f_{L_i}(l)$$

to the decoder. In this case, the superscript i is omitted in (8) above.

TABLE I
E_b/N_0 THRESHOLDS IN DECIBELS FOR ITERATIVE DECODING. 8-PSK MODULATION OVER AWGN CHANNEL

rate	C	C*	$\bar{\gamma}_1^*$	$\bar{\gamma}_2^*$	$\bar{\gamma}_3^*$	γ_1^*	γ_2^*	γ_3^*
1/6	-0.82	-0.80	0.93	4.24	7.47	0.91	4.37	7.55
1/5	-0.66	-0.62	0.92	3.84	6.92	0.88	3.95	6.99
1/4	-0.41	-0.36	1.03	3.39	6.28	0.97	3.49	6.34
1/3	0.00	0.12	1.53	2.94	5.51	1.44	3.01	5.56
1/2	0.86	1.28	5.41	2.76	4.65	5.31	2.78	4.69
2/3	1.76	2.75	6.25	3.82	5.17	6.20	3.83	5.18
3/4	2.23	3.66	6.76	4.53	5.62	6.73	4.54	5.63
4/5	2.51	4.30	7.11	5.04	5.98	7.10	5.05	5.98
5/6	2.70	4.75	7.39	5.43	6.26	7.38	5.43	6.27
6/7	2.84	5.11	7.61	5.74	6.50	7.60	5.74	6.50
7/8	2.94	5.41	7.80	5.99	6.70	7.79	5.99	6.70
8/9	3.02	5.66	7.95	6.21	6.87	7.95	6.21	6.87

D. Numerical Results

We trace the evolution of the bit LLR pdfs as the number of iterations ℓ increases, and determine σ^*, the maximum channel σ, such that for all $\sigma < \sigma^*$, $P_b \to 0$ as $\ell \to \infty$. We assume that $N^{2\ell}(v)$ is loop-free for all $v \in V(G)$, which implies independent messages [15].

In Tables I–IV, we report the E_b/N_0 thresholds γ^* above which $P_b \to 0$ as $\ell \to \infty$. The thresholds for i accumulate codes and m different channel densities are denoted γ_i^*, and the thresh-

TABLE II
E_b/N_0 THRESHOLDS IN DECIBELS FOR ITERATIVE DECODING. 8-PSK MODULATION OVER INDEPENDENT RAYLEIGH FADING CHANNEL

rate	C*	$\bar{\gamma}_1^*$	$\bar{\gamma}_2^*$	$\bar{\gamma}_3^*$	γ_1^*	γ_2^*	γ_3^*
1/6	-0.13	1.71	5.76	9.88	1.70	5.86	9.93
1/5	0.16	1.80	5.44	9.40	1.77	5.52	9.45
1/4	0.62	2.07	5.11	8.85	2.04	5.18	8.89
1/3	1.41	2.89	4.84	8.22	2.83	4.89	8.25
1/2	3.18	8.33	5.04	7.66	8.02	5.05	7.68
2/3	5.37	10.17	6.87	8.99	9.93	6.88	9.00
3/4	6.81	11.44	8.17	10.08	11.23	8.17	10.09
4/5	7.88	12.24	9.16	10.96	12.22	9.16	10.96
5/6	8.74	13.03	9.96	11.69	13.01	9.96	11.69
6/7	9.44	13.70	10.63	12.31	13.69	10.64	12.31
7/8	10.07	14.27	11.22	12.85	14.26	11.22	12.86
8/9	10.61	14.78	11.73	13.34	14.77	11.73	13.34

TABLE III
E_b/N_0 THRESHOLDS IN DECIBELS FOR ITERATIVE DECODING. 16-QAM OVER AWGN CHANNEL

rate	C	C*	$\bar{\gamma}_1^*$	$\bar{\gamma}_2^*$	$\bar{\gamma}_3^*$	γ_1^*	γ_2^*	γ_3^*
1/6	-0.55	-0.52	1.43	4.93	8.30	1.34	5.23	8.45
1/5	-0.33	-0.29	1.45	4.55	7.76	1.33	4.80	7.90
1/4	0.00	0.07	1.61	4.13	7.13	1.47	4.33	7.24
1/3	0.57	0.69	2.22	3.72	6.37	2.04	3.85	6.46
1/2	1.76	2.11	6.26	3.60	5.51	6.13	3.64	5.56
2/3	3.02	3.68	7.02	4.67	5.97	6.96	4.68	5.99
3/4	3.68	4.54	7.47	5.34	6.36	7.43	5.34	6.37
4/5	4.08	5.10	7.79	5.80	6.68	7.76	5.81	6.69
5/6	4.35	5.54	8.04	6.16	6.94	8.01	6.16	6.95
6/7	4.55	5.86	8.24	6.44	7.16	8.23	6.44	7.16
7/8	4.69	6.14	8.41	6.68	7.34	8.40	6.68	7.35
8/9	4.81	6.36	8.56	6.88	7.51	8.55	6.88	7.51

TABLE IV
E_b/N_0 THRESHOLDS IN DECIBELS FOR ITERATIVE DECODING. 16-QAM OVER INDEPENDENT RAYLEIGH FADING CHANNEL

rate	C*	$\bar{\gamma}_1^*$	$\bar{\gamma}_2^*$	$\bar{\gamma}_3^*$	γ_1^*	γ_2^*	γ_3^*
1/6	0.26	2.28	6.46	10.62	2.23	6.66	10.74
1/5	0.62	2.40	6.15	10.14	2.33	6.32	10.25
1/4	1.16	2.72	5.83	9.59	2.63	5.97	9.68
1/3	2.07	3.59	5.57	8.96	3.48	5.66	9.04
1/2	3.93	8.82	5.78	8.39	8.72	5.80	8.40
2/3	6.13	10.69	7.61	9.72	10.63	7.62	9.72
3/4	7.57	11.97	8.90	10.80	11.93	8.90	10.80
4/5	8.64	12.95	9.88	11.67	12.91	9.88	11.67
5/6	9.48	13.74	10.68	12.40	13.71	10.68	12.40
6/7	10.19	14.40	11.35	13.02	14.38	11.36	13.02
7/8	10.80	14.79	11.94	13.56	14.79	11.94	13.56
8/9	11.33	14.81	12.44	14.04	14.81	12.44	14.04

olds with one averaged channel density are denoted $\bar{\gamma}_i^*$. Rates $r_c < 1/2$ correspond to outer repeat codes, and rates $r_c > 1/2$ correspond to outer SPC codes. For $r_c = 1/2$, repeat and SPC codes are the same. For numerical reasons, we have defined the decoding threshold as the E_b/N_0 for which $P_b < 10^{-10}$ within 500 iterations.

In the tables, we also give the capacity for the channels under consideration. The 2-D channel capacity for an unconstrained discrete-time AWGN channel is $C = \log(1 + \text{SNR})$. For constrained input channels, such as PSK and QAM, and equiprobable signaling, the capacity for AWGN channels is given by [3]

$$C_A^* = \log M - \frac{1}{M}\sum_{k=0}^{M-1} \mathrm{E}_y \left[\log \sum_{j=0}^{M-1} e^{-\frac{\|y+x_k-x_j\|^2}{2\sigma^2}+\frac{\|y\|^2}{2\sigma^2}} \right]. \tag{9}$$

In (9), we use Monte-Carlo simulation to find the expected value of the random variable y, which is Gaussian distributed. For an independent Rayleigh fading channel, we extend (9) and get the constrained capacity as

$$C_R^* = \log M - \frac{1}{M}\sum_{k=0}^{M-1} \mathrm{E}_\rho \mathrm{E}_y \left[\log \sum_{j=0}^{M-1} e^{-\frac{\|y+\rho x_k-\rho x_j\|^2}{2\sigma^2}+\frac{\|y\|^2}{2\sigma^2}} \right]. \tag{10}$$

Again, we use Monte-Carlo simulation to evaluate the double expectation in (10), where the fading amplitude ρ is Rayleigh distributed. In Tables I–IV, C and C^* denote the minimum E_b/N_0 needed to achieve the corresponding rate for an unconstrained and a constrained channel, respectively.

In the tables, we notice that for an outer repeat code, a single inner accumulate code gives the lowest threshold. For an outer SPC, the lowest threshold is achieved with two accumulate codes. We know from the coding theorems in [14] that for PA^1 codes, $P_b \to 0$ as $N \to \infty$, but P_W does not. This implies that the number of bit errors in a word $n_e \neq 0$, but n_e grows slower than N (or stays constant), so $P_b = n_e/N \to 0$ as $N \to \infty$. Therefore, the thresholds for PA^1 codes are really thresholds above which n_e grows slower than $P_b N$. Note that for PA^1 codes, the value of the E_b/N_0 thresholds in Tables I–IV depend on the choice $P_b < 10^{-10}$. In Section IV, we derive a stability condition which provides a lower bound on the E_b/N_0 threshold. For all other codes, the computed thresholds are insensitive to the choice of the P_b threshold.

We can consider a PA^2 code as an outer PA^1 code and one inner accumulate code. The outer code has a $d_{\min}^{(o)} \geqslant 3$ with a probability approaching 1 as $N \to \infty$. Hence, for the concatenation PA^2, $P_W \to 0$ with high probability as $N \to \infty$.

For PA^2 codes of rate $r_c \geqslant 2/3$, the thresholds are about 1 dB away from the constrained capacity for AWGN channels, and about 1.5 dB away for Rayleigh fading. For RA^1 codes, the gap to the constrained capacity is between 1.5 and 2 dB for both AWGN and Rayleigh fading. All codes under consideration are regular. By using irregular RA codes [25], [26] or irregular PA codes, we expect to reduce the gap to capacity.

We also notice that the differences between the thresholds computed with m different channel densities and the thresholds computed with one averaged density are small, in particular for higher rates. Using one averaged channel density simplifies the threshold computations.

IV. STABILITY CONDITION

As noted in the previous section, the DE thresholds for PA^1 codes depend on the choice of target P_b. To avoid this, we now

derive a stability condition similar to the one in [16]. We assume that decoding has progressed successfully to a point where only a small fraction of bits remain in error, that is, the LLR density is close to a unit impulse. The stability condition gives the minimum channel SNR required to ensure that the LLR density becomes a unit impulse at positive infinity as the number of decoding iterations goes to infinity. If the messages in the DE, the pdfs, are symmetric in the sense of (5), then a unit impulse at $+\infty$ implies that $f(l) = 0$ for all $l < \infty$, and hence, the bit-error probability (BEP) $P_b = \int_{-\infty}^{0} f(l)dl = 0$.

Let a message consist of a discrete part at positive infinity, δ_∞, and a continuous general density $Q(l) \neq \delta_\infty$, $\int_{-\infty}^{\infty} Q(l)dl = 1$, that is

$$\mathsf{m}(l) = [(1-\varepsilon)\delta_\infty + \varepsilon Q(l)]$$

where $1-\varepsilon$ is the fraction of the pdf at $+\infty$, and ε is the fraction in the continuous part. We assume that a sufficient number of iterations have been performed so ε is small. For the decoder to have a fixed point, i.e., to be stable, at $P_b = 0$, the continuous part Q of the message must vanish as the number of iterations goes to infinity. We wish to find the lowest E_b/N_0 for which Q vanishes.

We consider a rate $r_c = (n-1)/n$ parity code and simplify the notation by incorporating the state information in the bit node. The case of one averaged channel density was treated in [27], and here we also analyze the case of multiple channel densities.[3]

A. Averaged Channel Density

When the channel interleaver Ψ is present, we average the pdfs from the channel. This simplifies the analysis, since all messages from a given node will be the same. A fragment of the corresponding decoding graph is shown in Fig. 3(a).

In general, the operations performed at the nodes can be any two operations, as long as they are associative and commutative, and preserve the symmetry of the pdfs. In this case, the operations are $\otimes$ and $\boxplus$, as previously described, which are associative and commutative. It was shown in [16] that they preserve the pdf symmetry.

From the graph in Fig. 3(a), we get the following relationships between the messages:

$$\begin{aligned} a &= b \otimes f_L \\ c &= a \boxplus a \\ d &= c^{\boxplus(n-1)} \\ b &= a \boxplus d \\ \mathsf{m}_d &= c^{\boxplus n}. \end{aligned}$$

The update equation for the message a and the decision message m_d in terms of the message a are

$$a = a^{\boxplus(2n-1)} \otimes f_L \tag{11}$$

$$\mathsf{m}_d = a^{\boxplus 2n}. \tag{12}$$

[3]The notation for the single-density case is different in this paper in order to be compatible with the notation for the multiple-density case.

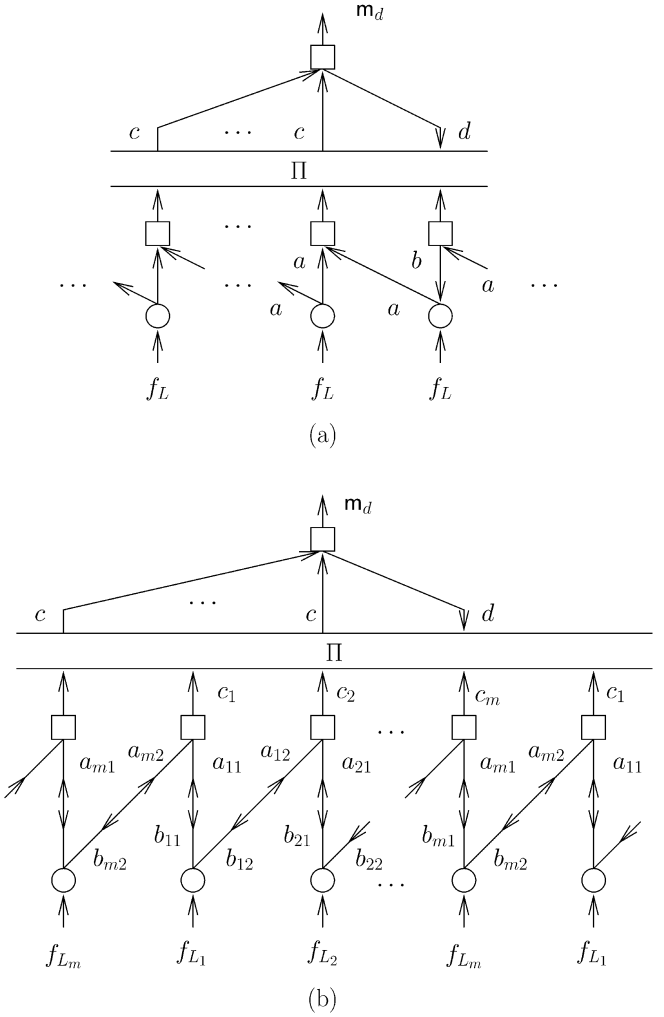

Fig. 3. (a) Fragment of the simplified decoding graph with a common, averaged channel density. The one-input, one-output bit node between the accumulate code and the SPC is omitted for clarity. (b) Fragment of decoding graph with multiple input densities.

The probability of bit error is

$$P_b = \int_{-\infty}^{0} \mathsf{m}_d(l)dl.$$

For P_b to go to zero, the continuous part of the message m_d, Q_{m_d}, must vanish on the interval $(-\infty, 0]$ as the number of iterations (ℓ) goes to infinity. Since m_d is symmetric, this implies that P_b goes to zero if and only if $\mathsf{m}_d(l) \to \delta_\infty$ as the number of iterations goes to infinity.

In order to expand (12), note that $Q \boxplus \delta_\infty = Q$, $\delta_\infty \boxplus \delta_\infty = \delta_\infty$, and $Q \otimes \delta_\infty = \delta_\infty$. Expanding the update equation for message a as given in (11)

$$\begin{aligned} a &= a^{\boxplus(2n-1)} \otimes f_L \\ &= \left[(1-\varepsilon_a)\delta_\infty + \varepsilon_a Q_a\right]^{\boxplus(2n-1)} \otimes f_L \\ &= \Big[(1-\varepsilon_a)^{(2n-1)}\delta_\infty^{\boxplus(2n-1)} + (2n-1)(1-\varepsilon_a)^{(2n-1)-1} \\ &\qquad \times \varepsilon_a Q_a \boxplus \delta_\infty^{\boxplus(2n-1)-1} + O\left(\varepsilon_a^2\right)\Big] \otimes f_L. \end{aligned}$$

Assuming ε_a small, we neglect ε_a^2 and higher terms, yielding

$$a = [(1-(2n-1)\varepsilon_a)\,\delta_\infty + (2n-1)\varepsilon_a Q_a] \otimes f_L. \quad (13)$$

After ℓ iterations, the update equation is

$$a = \left(1-(2n-1)^\ell \varepsilon_a\right)\delta_\infty + (2n-1)^\ell \varepsilon_a Q_a \otimes f_L^{\otimes \ell}$$

and the decision message m_d after ℓ iterations is

$$\mathsf{m}_d = \left(1-2n(2n-1)^\ell \varepsilon_a\right)\delta_\infty + 2n(2n-1)^\ell \varepsilon_a Q_a \otimes f_L^{\otimes \ell}.$$

The BEP P_b is given by

$$P_b = \int_{-\infty}^{0} 2n(2n-1)^\ell \varepsilon_a Q_a(l) \otimes f_L(l)^{\otimes \ell} dl. \quad (14)$$

Since f_L is a symmetric density and $Q \neq \delta_\infty$, (14) goes to zero only if

$$\int_{-\infty}^{0} 2n(2n-1)^\ell f_L(l)^{\otimes \ell} dl \quad (15)$$

goes to zero as $\ell \to \infty$. Using Chernoff's large deviation result, as in [16], the condition in (15) can be restated as saying that

$$(2n-1)^\ell \int_{-\infty}^{0} f_L(l)^{\otimes \ell} dl \quad (16)$$

must go to zero as $\ell \to \infty$. By defining the parameter r as

$$r := -\ln \int_{-\infty}^{\infty} f_L(l) e^{-\frac{l}{2}} dl$$

the condition in (16) reduces to

$$2n-1 < e^r. \quad (17)$$

Finding the lowest channel SNR such that (17) is met gives us a E_b/N_0 threshold $\bar{\gamma}_{\mathrm{sc}}$, below which $P_b > 0$. That is, the stability condition is a necessary condition and provides a lower bound on the E_b/N_0 threshold. It is stated without proof in [16] that the condition is both necessary and sufficient. The stability condition would, therefore, also provide an E_b/N_0 threshold above which $P_b \to 0$ as $\ell \to \infty$.

B. Multiple Channel Densities

In this section, we derive a stability condition where we consider m different densities from the channel. We now need to keep track of different messages in the decoding of the accumulate code. Because of the interleaver between the accumulate code and the SPC, the messages are averaged before decoding the SPC. Hence, the messages in and out of the SPC nodes are all the same. A fragment of the decoding graph is shown in Fig. 3(b).

In the previous section, the message a and b were the same for all nodes. Here we need to account for multiple channel densities and index the messages. The a messages become $a_{ij} = [(1-\varepsilon_{ij})\delta_\infty + \varepsilon_{ij} Q_{ij}]$, where $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant 2$. Again, we express the other messages in terms of a_{ij}. All index additions and subtractions are performed modulo m. From Fig. 3(b), we get

$$a_{i1} = \left(a_{(i+1)1} \boxplus d\right) \otimes f_{L_i}$$
$$a_{i2} = \left(a_{(i-1)2} \boxplus d\right) \otimes f_{L_i}.$$

At the input of the interleaver, we have

$$c_i = a_{i1} \boxplus a_{(i-1)2}$$

and at the output of the interleaver

$$\begin{aligned} c &= \frac{1}{m}\sum_{i=1}^{m} c_i \\ &= \frac{1}{m}\left[(m-\varepsilon_{11}-\varepsilon_{12}-\ldots-\varepsilon_{m2})\delta_\infty \right. \\ &\qquad \left. +\varepsilon_{11}Q_{11}+\varepsilon_{12}Q_{12}+\ldots+\varepsilon_{m2}Q_{m2}\right]. \end{aligned}$$

Define the averages $\varepsilon := (1/2m)\sum_{ij}\varepsilon_{ij}$ and $Q := (\sum_{ij}\varepsilon_{ij}Q_{ij}/\sum_{ij}\varepsilon_{ij})$. Then we have

$$c = \left[(1-2\varepsilon)\delta_\infty + 2\varepsilon Q\right].$$

The output from the SPC is

$$d = c^{\boxplus(n-1)} = \left[(1-2(n-1)\varepsilon)\,\delta_\infty + 2(n-1)\varepsilon Q\right]$$

and the decision message is

$$\mathsf{m}_d = c^{\boxplus n} = \left[(1-2n\varepsilon)\delta_\infty + 2n\varepsilon Q\right].$$

The update equations for $a_{11}, \ldots, a_{m2}$ are

$$\begin{aligned} a_{i1} &= \left[\left(1-\varepsilon_{(i+1)1}-2(n-1)\varepsilon\right)\delta_\infty \right. \\ &\qquad \left. +\left(\varepsilon_{(i+1)1}Q_{(i+1)1}+2(n-1)\varepsilon Q\right)\otimes f_{L_i}\right] \\ a_{i2} &= \left[\left(1-\varepsilon_{(i-1)2}-2(n-1)\varepsilon\right)\delta_\infty \right. \\ &\qquad \left. +\left(\varepsilon_{(i-1)2}Q_{(i-1)2}+2(n-1)\varepsilon Q\right)\otimes f_{L_i}\right]. \end{aligned} \quad (18)$$

It is interesting to compare these expressions with (13), the update equation for a in the case of a single averaged channel density. Here, Q is an average over the Q_{ij}, and if all Q_{ij} are equal, the expressions in (18) are identical to (13). Instead of averaging the channel densities before the decoder, the averaging takes place within the decoder.

TABLE V
STABILITY CONDITION E_b/N_0 THRESHOLDS IN DECIBELS FOR 8-PSK AND 16-QAM

	8-PSK				16-QAM			
	AWGN		Fading		AWGN		Fading	
rate	$\bar{\gamma}_{\rm sc}$	$\gamma_{\rm sc}$	$\bar{\gamma}_{\rm sc}$	$\gamma_{\rm sc}$	$\bar{\gamma}_{\rm sc}$	$\gamma_{\rm sc}$	$\bar{\gamma}_{\rm sc}$	$\gamma_{\rm sc}$
1/2	5.28	5.19	7.89	7.83	6.15	6.01	8.63	8.53
2/3	6.18	6.14	9.82	9.79	6.96	6.90	11.10	11.06
3/4	6.71	6.68	11.12	11.10	7.57	7.55	13.00	12.97
4/5	7.07	7.06	12.11	12.10	8.05	8.04	14.08	14.08
5/6	7.35	7.34	12.70	12.70	8.40	8.39	14.35	14.35
6/7	7.58	7.57	12.84	12.84	8.68	8.68	14.55	14.55
7/8	7.77	7.76	12.94	12.94	8.92	8.91	14.69	14.69
8/9	7.93	7.92	13.02	13.02	9.11	9.11	14.81	14.81

The stability condition for m densities is somewhat more involved than the corresponding criterion for an averaged density. Let $\mathbf{Q}$ be a vector of densities Q_{ij}, such that the first element of $\mathbf{Q}$ is the sum of the two densities Q_{ij} that are convolved with f_{L_1} in the update equations for a_{ij} (18). Let $\mathbf{f}_L$ be a column vector of the channel densities $f_{L_1}, \ldots, f_{L_m}$ and f_L be the averaged channel density.

The message c after ℓ iterations can then be expressed as

$$c = \left[(1-2\varepsilon)\delta_\infty + \varepsilon Q^\ell\right]$$

and the update equation for the continuous density Q is given by

$$Q^\ell = 2(n-1)Q^{\ell-1} \otimes f_L + \mathbf{Q}^\ell \otimes \mathbf{f}_L$$

where $\mathbf{Q}^\ell$ is the vector $\mathbf{Q}$ after ℓ iterations. The BEP as $\ell \to \infty$ is then

$$\begin{aligned} P_b &= \int_{-\infty}^{0} \lim_{\ell\to\infty} c(l)^{\boxplus n} dl \\ &= \int_{-\infty}^{0} \lim_{\ell\to\infty} 2n \sum_{i=1}^{\ell} [2(n-1)]^i \, \mathbf{Q}^{\ell-i} \otimes \mathbf{f}_L \otimes f_L^{\otimes i} dl. \end{aligned} \quad (19)$$

If the m channel densities are all equal, then $\mathbf{Q}(l) \otimes \mathbf{f}_L(l) = Q(l) \otimes f_L(l)$, and (19) collapses to (14).

Since the expressions for the single density and multiple densities are almost the same, we do not expect any large differences in the threshold values. Indeed, in Table V, we see only small differences between the thresholds obtained from the two methods. The threshold for an averaged density is denoted $\bar{\gamma}_{\rm sc}$, and the threshold for m densities is denoted $\gamma_{\rm sc}$.

V. SIMULATION RESULTS

In Fig. 4, we compare simulated bit-error rate (BER) performance of an RA code to the DE thresholds for 8-PSK modulation over AWGN. The code rate is 1/3 and the blocklength $N = 10^5$. The DE threshold is 1.44 dB, and is indicated with a vertical dash-dotted line. The simulated performance corresponds well to the threshold.

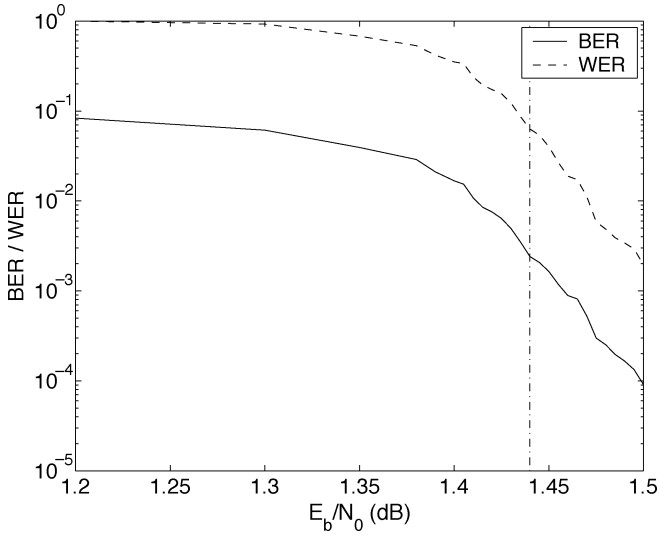

Fig. 4. Comparison of simulations to the DE threshold for an $r_c = 1/3$ RA code with blocklength $N = 10^5$. The vertical, dot-dashed line is the DE threshold.

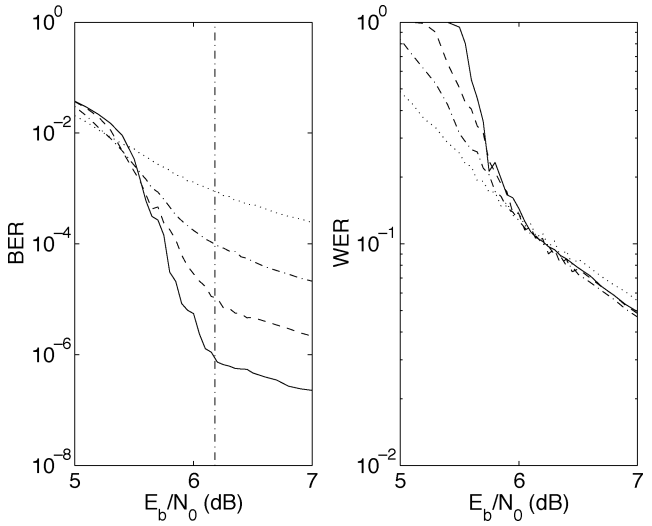

Fig. 5. BER and WER as a function of the blocklength N. For the dotted line, $N = 10^3$, for the dot-dashed line, $N = 10^4$, dashed line, $N = 10^5$, and for the solid line, $N = 10^6$. The vertical, dot-dashed line is the stability condition threshold.

In Fig. 5, we show the BER and WER for rate-2/3 PA codes of different blocklengths for 8-PSK modulation over AWGN. We compare the BER with the stability condition threshold at 6.18 dB. As the blocklength increases, the BER decreases, but the WER does not depend on the blocklength. This is consistent with the coding theorems in [14].

In Fig. 6, we show the BER and WER for rate-2/3 PA^2 and PA^3 codes, blocklength $N = 10^5$, 8-PSK modulation over an AWGN channel. As the SNR exceeds the DE thresholds, both the BER and WER decrease rapidly.

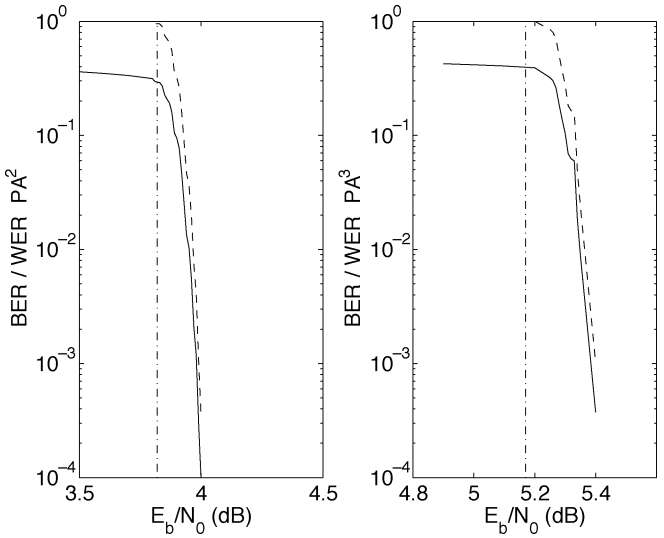

Fig. 6. BER (solid lines) and WER (dashed lines) for PA^2 and PA^3 codes, left and right plots, resp., with 8-PSK modulation over an AWGN channel. The blocklength is $N = 10^5$, and the dot-dashed lines are the DE thresholds.

VI. Conclusions

We have analyzed an SCTCM system with one or more inner accumulate codes. We have used DE to compute E_b/N_0 thresholds for an iterative message-passing decoder. For PA^1 codes, we have also computed lower bounds on the E_b/N_0 thresholds using a stability condition.

We have devised a method to compute the LLR pdfs for higher order constellations, and shown that these pdfs are symmetric. Furthermore, the m different LLR pdfs of a higher order constellation can be averaged without significantly changing the computed DE and stability condition thresholds.

For PA^2 codes, which have the lowest thresholds of the PA^i codes, and 8-PSK modulation over AWGN channels, the thresholds are about 1 dB away from the constrained capacity for $r_c = 2/3$, and closer to capacity for higher rates.

Appendix
Symmetry of Log-Likelihood PDFs

Let the random variable L be an LLR and the value l an outcome of L. In this appendix, we show that the pdf $f_L(l)$ for the LLR is symmetric in the sense of [16], i.e., that

$$f_L(-l) = e^{-l} f_L(l). \tag{20}$$

As in Section III-A, $\mathcal{S}$ is a signal constellation of size $|\mathcal{S}| = M = 2^m$, with a binary labeling of m bits, $\mathcal{S}_0$ is the subset of $\mathcal{S}$ where the ith bit in the label is zero, and $\mathcal{S}_1 = \mathcal{S} \setminus \mathcal{S}_0$.

The log-likelihood functions $g_0(r)$ and $g_1(r)$ are defined in (2). If the subscript 0 or 1 is omitted, the choice of function $g_0(r)$ or $g_1(r)$ depends on the transmitted symbol s. For a given value r, the only difference between $g_0(r)$ and $g_1(r)$ in (2) is the sign.

The received value r is an outcome of the random variable R, whose mean depends on the transmitted symbol s and variance depends on the channel SNR. The function $f_R(r|s)$ is the pdf for the received value r, given that the symbol s was transmitted. In two dimensions, let $r = (x, y)$, $s = (s_x, s_y)$, and the functions in (2) become $g_0(x, y)$ and $g_1(x, y)$. For independent Gaussian random variables X and Y, we get the pdf of the received values x and y, conditioned on the transmitted symbol s, as

$$f_{XY}(x, y|s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-s_x)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-s_y)^2}{2\sigma^2}}.$$

Assuming all symbols are transmitted equally likely, the unconditional pdf is given by (3). The LLR L is a function of two random variables, $L = g(X, Y)$, and the pdf of the LLR, $f_L(l)$, is calculated by (4). If, for fixed values l and y, the equation $x = g^{-1}(l, y)$ has several solutions, (4) should be integrated over all such solutions.

A. Proof for 2-D Constellations

We are now ready to verify (20). Define, for given values of l and y, the sets

$$\begin{aligned} x_0^+ &:= \{x \in \mathbb{R} : g_0(x, y) = +l\} \\ x_0^- &:= \{x \in \mathbb{R} : g_0(x, y) = -l\} \\ x_1^+ &:= \{x \in \mathbb{R} : g_1(x, y) = +l\} \\ x_1^- &:= \{x \in \mathbb{R} : g_1(x, y) = -l\} \end{aligned} \tag{21}$$

where sets may be empty. By definition, $g_0(x, y) = -g_1(x, y)$, so $x_0^+ = x_1^-$ and $x_0^- = x_1^+$.

Inserting first (3) into (4) and then (4) into (20) yields

$$e^{-l} f_L(l) = e^{-l} \int_{-\infty}^{\infty} |J|^{-1} \frac{1}{M} \sum_{j=1}^{M} f_{XY}(x, y|s_j) \Big|_{x=g^{-1}(l,y)} dy. \tag{22}$$

Split the summation into sums over $\mathcal{S}_0$ and $\mathcal{S}_1$, and let J_0 and J_1 denote the Jacobian evaluated using $g_0(x, y)$ and $g_1(x, y)$, respectively. The right-hand side of (22) becomes

$$e^{-l} \int_{-\infty}^{\infty} \left(|J_0|^{-1} \frac{1}{M} \sum_{j \in \eta_0} f_{XY}(x, y|s_j) \Big|_{x=g_0^{-1}(l,y)} + |J_1|^{-1} \frac{1}{M} \sum_{j \in \eta_1} f_{XY}(x, y|s_j) \Big|_{x=g_1^{-1}(l,y)} \right) dy. \tag{23}$$

Recall that we sum (integrate) over all probability contributions for which $x = g^{-1}(l, y)$. Bring e^{-l} inside the integration to get

$$\frac{1}{M} \int_{-\infty}^{\infty} \left(|J_0|^{-1} e^{-l} \sum_{j \in \eta_0} f_{XY}(x, y|s_j) \Big|_{x=g_0^{-1}(l,y)} + |J_1|^{-1} e^{-l} \sum_{j \in \eta_1} f_{XY}(x, y|s_j) \Big|_{x=g_1^{-1}(l,y)} \right) dy. \tag{24}$$

By the definitions in (21), $g_1(x, y) = -l$ for $x \in x_0^+$, and $g_0(x, y) = -l$ for $x \in x_1^+$, so $e^{-l} = e^{g_1(x,y)}$ for any $x \in x_0^+$ and $e^{-l} = e^{g_0(x,y)}$ for any $x \in x_1^+$. Then (24) can be written

$$\frac{1}{M} \int_{-\infty}^{\infty} \left(|J_0|^{-1} e^{g_1(x,y)} \sum_{j \in \eta_0} f_{XY}(x, y|s_j) \Big|_{x=g_0^{-1}(l,y)} + |J_1|^{-1} e^{g_0(x,y)} \sum_{j \in \eta_1} f_{XY}(x, y|s_j) \Big|_{x=g_1^{-1}(l,y)} \right) dy. \tag{25}$$

Using the definitions of g_0 and g_1, we get

$$\frac{1}{M}\int_{-\infty}^{\infty}\left(|J_0|^{-1}\frac{\sum_{j\in\eta_1}f_{XY}(x,y|s_j)}{\sum_{j\in\eta_0}f_{XY}(x,y|s_j)}\times\sum_{j\in\eta_0}f_{XY}(x,y|s_j)\Big|_{x=g_0^{-1}(l,y)}+|J_1|^{-1}\frac{\sum_{j\in\eta_0}f_{XY}(x,y|s_j)}{\sum_{j\in\eta_1}f_{XY}(x,y|s_j)}\times\sum_{j\in\eta_1}f_{XY}(x,y|s_j)\Big|_{x=g_1^{-1}(l,y)}\right)dy. \quad (26)$$

Cancel terms to reduce (26) to

$$\frac{1}{M}\int_{-\infty}^{\infty}\left(|J_0|^{-1}\sum_{j\in\eta_1}f_{XY}(x,y|s_j)\Big|_{x=g_0^{-1}(l,y)}+|J_1|^{-1}\sum_{j\in\eta_0}f_{XY}(x,y|s_j)\Big|_{x=g_1^{-1}(l,y)}\right)dy. \quad (27)$$

Since $x_0^+ = x_1^-$ and $x_0^- = x_1^+$, and $|J_0| = |J_1|$, we get

$$\begin{aligned}&\frac{1}{M}\int_{-\infty}^{\infty}\left(|J_1|^{-1}\sum_{j\in\eta_1}f_{XY}(x,y|s_j)\Big|_{x=g_1^{-1}(-l,y)}+|J_0|^{-1}\sum_{j\in\eta_0}f_{XY}(x,y|s_j)\right)\Bigg|_{x=g_0^{-1}(-l,y)}dy\\&=\frac{1}{M}\int_{-\infty}^{\infty}\left(|J_1|^{-1}\sum_{j\in\eta_1}f_{XY}(x,y|s_j)\Big|_{x=g_1^{-1}(-l,y)}+|J_0|^{-1}\sum_{j\in\eta_0}f_{XY}(x,y|s_j)\Big|_{x=g_0^{-1}(-l,y)}\right)dy\\&=\int_{-\infty}^{\infty}\left(|J|^{-1}\frac{1}{M}\sum_{j=1}^{M}f_{XY}(x,y|s_j)\Big|_{x=g^{-1}(-l,y)}\right)dy\\&=f_L(-l).\end{aligned}$$

■

B. Proof for 1-D Constellations

In the 1-D case, the LLR functions g_0 and g_1 are functions of one random variable. The Jacobian determinants become the derivatives of $g_0(x)$ and $g_1(x)$ with respect to x. The proof of symmetry follows the proof for the 2-D case.

ACKNOWLEDGMENT

The authors wish to thank the reviewers for their valuable comments and suggestions on improving this paper.

REFERENCES

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

[2] 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, pp. 909–926, May 1998.

[3] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.

[4] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Parallel concatenated trellis coded modulation," in *Proc. IEEE Int. Conf. Commun.*, vol. 2, Dallas, TX, Jun. 1996, pp. 974–978.

[5] P. Robertson and T. Wörz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 201–218, Feb. 1998.

[6] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenated trellis coded modulation with iterative decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Ulm, Germany, Jun.–Jul. 1997, p. 8.

[7] D. Divsalar, S. Dolinar, and F. Pollara, "Serial concatenated trellis coded modulation with rate-1 inner code," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, CA, Nov.–Dec. 2000, pp. 777–782.

[8] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for "turbo-like" codes," in *Proc. 36th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 1998, pp. 201–210.

[9] H. D. Pfister and P. H. Siegel, "The serial concatenation of rate-1 codes through uniform random interleavers," in *Proc. 37th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 1999, pp. 260–269.

[10] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.

[11] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding using soft feedback," *Electron. Lett.*, vol. 34, pp. 942–943, May 1998.

[12] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Commun.*, vol. 50, pp. 1250–1257, Aug. 2002.

[13] H. M. Tullberg and P. H. Siegel, "Serial concatenated trellis coded modulation with inner rate-1 accumulate code," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, San Antonio, TX, Nov. 2001, pp. 936–940.

[14] ——, "Serial concatenated TCM with an inner accumulate code—Part I: Maximum-likelihood analysis," *IEEE Trans. Commun.*, vol. 53, pp. 64–73, Jan. 2005.

[15] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, pp. 599–618, Feb. 2001.

[16] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, pp. 619–637, Feb. 2001.

[17] 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.

[18] D. Divsalar and F. Pollara, "Turbo codes for PCS applications," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, Jun. 1995, pp. 54–59.

[19] T. J. Richardson and R. L. Urbanke, "An introduction to the analysis of iterative coding systems," in *IMA 1999 Summer Program: Codes, Systems and Graphical Models*. New York: Springer-Verlag, 2001, pp. 1–37.

[20] J. Hou, P. H. Siegel, L. B. Milstein, and H. D. Pfister, "Multilevel coding with low-density parity-check component codes," in *Proc. IEEE Global Telecommun. Conf.*, San Antonio, TX, Nov. 2001, pp. 1016–1020.

[21] C. W. Helstrom, *Probability and Stochastic Processes for Engineers*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1991.

[22] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[23] J. Hagenauer. The turbo principle: Tutorial introduction and state of the art. presented at *Int. Symp. Turbo Codes, Related Topics*. [Online]. Available: http://www.lnt.e-technik.tu-muenchen.de/veroeffentlichungen/tut.ps.gz

[24] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[25] H. Jin, A. Khandekar, and R. J. McEliece, "Irregular repeat-accumulate codes," in *Proc. Int. Symp. Turbo Codes, Related Topics*, Brest, France, Sep. 2000, pp. 1–8.
[26] H. Jin, "Analysis and design of turbo-like codes," Ph.D. dissertation, Calif. Inst. of Technol., Pasadena, CA, May 2001.
[27] H. M. Tullberg and P. H. Siegel, "Serial concatenated trellis coded modulation with inner rate-1 accumulate code," in *Proc. IEEE Veh. Technol. Conf.*, vol. 4, Atlantic City, NJ, Oct. 2001, pp. 2333–2337.

Hugo M. Tullberg (S'98–M'03) 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 at San Diego, La Jolla, in 2002.

He is currently a Senior Scientist with the Swedish Defence Research Agency (FOI), Linköping, Sweden. His research interests include information theory, coding theory, communications, and graph-based systems. He is also working on problems in ad hoc networking.

Dr. Tullberg is a member of the IEEE Communications, Information Theory, and Vehicular Technology Societies.

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 (MIT), Cambridge.

He held a Chaim Weizmann Postdoctoral Fellowship at the Courant Institute, New York University. He was with the IBM Research Division in San Jose, CA, 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 California Institute of Telecommunications and Information Technology, the Center for Wireless Communications, and the Center for Magnetic Recording Research where he currently serves as Director. His primary research interests lie in the areas of information theory and communications, particularly coding and modulation techniques, with applications to digital data storage and transmission.

Prof. Siegel 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, served the same Transactions as Associate Editor for Coding Techniques from 1992 to 1995, and as Editor-in-Chief from 2001 to 2004. He was also Co-Guest Editor of the May/September 2001 two-part issue on The Turbo Principle: From Theory to Practice of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He was co-recipient, with R. Karabed, of the 1992 IEEE Information Theory Society Paper Award and shared the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award with B. Marcus and J. K. Wolf. He holds 17 patents in the area of coding and detection, and was named a Master Inventor at IBM Research in 1994. He is also a member of Phi Beta Kappa.