Design of Repeat–Accumulate Codes for Iterative Detection and Decoding

Stephan ten Brink and Gerhard Kramer, *Member, IEEE*

***Abstract*—Extrinsic information transfer (EXIT) charts are used to design systematic and nonsystematic repeat-accumulate (RA) codes for iterative detection and decoding. The convergence problems of nonsystematic RA codes are solved by introducing a biregular, or doped, layer of check nodes. As examples, such nonsystematic codes are designed for multi-input/multi-output (MIMO) fading channels and are shown to operate close to capacity.**

***Index Terms*—Detection, iterative decoding, mutual information, repeat-accumulate codes.**

I. Introduction

REPEAT-ACCUMULATE (RA) codes [1] are a competitive alternative to parallel concatenated (PC or turbo) [2] and low-density parity-check (LDPC) codes [3], [4]. Irregular RA codes, as devised by Jin *et al.* [5], consist of an outer code that is a mixture of repetition codes and an inner code that consists of parity checks and an accumulator (a differential encoder). We revisit the design of RA codes by using extrinsic information transfer (EXIT) charts [6]. Our approach follows closely that described for LDPC codes in [7], but several new and interesting design issues arise because of the accumulator.

Perhaps the main advantage of irregular RA codes over irregular LDPC codes [8]–[11] is that the complexity of encoding is linear in the code length. On the other hand, an iterative LDPC decoder can be fully "parallelized" while an *a posteriori* probability (APP) iterative RA decoder cannot because of the accumulator.[1] However, both RA and LDPC codes enjoy an advantage over turbo codes in that they offer a more flexible code structure, with many parameters (degrees) that can be adjusted according to a particular design criterion. In this context, we point out that irregular turbo codes [12], [13] are closely related to irregular RA codes, the main difference being that a trellis with memory two or more replaces the accumulator.

Two classes of irregular RA codes were introduced in [5]: systematic and nonsystematic. Both classes are *check-regular* in the sense that the encoder has a layer of check nodes where each check node takes as inputs exactly d_c edges from an interleaver. *Systematic* RA codes are not serially concatenated codes because both the inner and outer decoders receive channel information. In contrast, *nonsystematic* RA codes are serially concatenated codes, i.e., only the inner decoder receives channel information. We find that this property makes the nonsystematic codes simpler to design when combined with a modulator and detector. However, as noted in [5], nonsystematic RA decoders require $d_c > 1$ to approach capacity, yet do not even begin to converge for such cases. We show how to overcome this problem by making the check node layer *biregular*, i.e., the check nodes have either degree 1 or d_c. This trick is a variant of *code doping* [14], and it enables a successful start to the iterative decoding.

We remark that our design methodology can be applied to any detector, e.g., noncoherent detectors or equalizers for intersymbol interference channels. We here concentrate on multi-input, multi-output (MIMO) detectors because they will likely play a prominent role in future wireless and wireline communication systems.

This paper is organized as follows. In Section II, we show how to compute EXIT curves for RA codes used with binary phase shift keying (BPSK) on additive white Gaussian noise (AWGN) channels. We design both systematic and novel nonsystematic RA codes that closely approach capacity. Section III extends the AWGN design methods to combined detection and decoding for MIMO fading channels. Section IV summarizes our results.

II. EXIT Chart for RA Codes

A. Systematic and Nonsystematic RA Encoders

We follow [5] and distinguish between *systematic* and *nonsystematic* RA codes. The encoder and decoder for a *systematic* RA code of rate $R = k/n$ are both depicted in Fig. 1. The encoder consists of four main parts: k repetition codes with variable rates, an interleaver, $n - k$ single parity check codes, and a differential encoder or *accumulator*. The repetition codes are represented as *variable nodes*, the interleaver as an edge interleaver, the single parity check codes as *check nodes*, and the accumulator as a chain of check nodes.

Suppose the rate of the ith repetition code is $1/d_v^{(i)}$. We adopt the terminology used for LDPC codes and define the node *degrees* with respect to the edge interleaver. Thus, we say that the ith variable node degree is $d_v^{(i)}$, even though this node has $d_v^{(i)} + 1$ edges touching it. Similarly, let the ith check node degree be the number $d_c^{(i)}$ of edges incident to this node from the edge interleaver. The rate of the ith single parity check code is therefore $d_c^{(i)}/(d_c^{(i)} + 1)$.

Manuscript received December 13, 2002; revised May 12, 2003. The associate editor coordinating the review of this paper and approving it for publication was Prof. Brian Hughes.

S. ten Brink was with Bell Laboratories, Lucent Technologies, Crawford Hill, NJ USA. He is now with Realtek, Irvine, CA 92618 USA (e-mail: stenbrink@realtek-us.com).

G. Kramer is with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07928 USA (e–mail: gkr@bell-labs.com).

Digital Object Identifier 10.1109/TSP.2003.818250

[1]By an "APP iterative RA decoder," we mean performing a forward and backward pass through the accumulator trellis every iteration. Another approach is to parallelize the decoder by updating trellis sections only locally. However, decoding will then require more iterations.

1053-587X/03$17.00 © 2003 IEEE

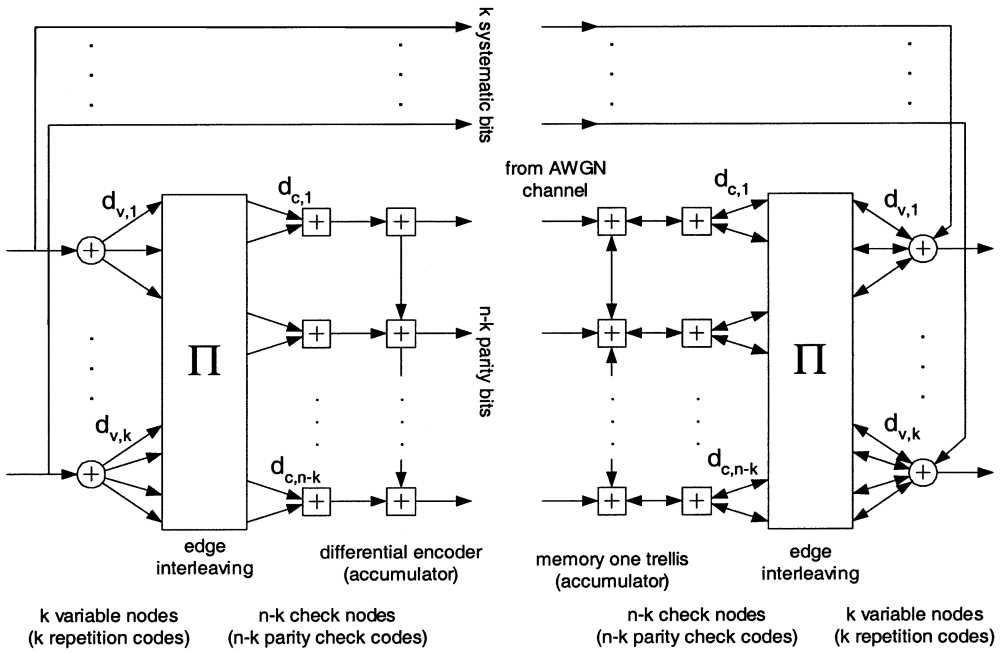


Fig. 1. Encoder and decoder for a systematic RA code.

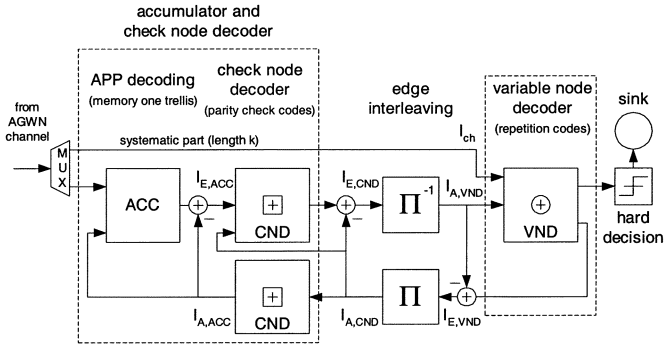


Fig. 2. Decoder for a systematic RA code.

For *nonsystematic* RA codes, the information bits are not transmitted. We therefore now have n inner single parity check codes and n codeword bits coming out of the accumulator.

B. Systematic and Nonsystematic RA Decoders

Consider first the decoder for a *systematic* RA code, as depicted in Fig. 2. The inner decoder takes the $n-k$ channel reliability values (log-likelihood ratio values or L-values [15]) for the parity bits and performs an APP decoding over the memory-one trellis of the accumulator. The resulting output L-values are forwarded to the inner check node decoder (CND) comprising the $n-k$ check nodes. The CND generates L-values that are forwarded through the edge interleaver to the outer variable node decoder (VND) comprising k variable nodes. The VND uses these L-values and the k systematic L-values to generate new L-values that are fed back through the edge interleaver to the CND. Finally, one decoder iteration is completed when the CND generates L-values that it feeds back to the accumulator decoder (ACC). We will often consider the ACC and the CND as a single decoding unit that we refer to as the ACC&CND-decoder.

For *nonsystematic* RA codes, the accumulator length is n, and there are n check nodes. Of course, there is now no systematic input to the outer variable node decoder.

C. EXIT Curve of the Outer Variable Node Decoder

We begin with the VND, whose EXIT curve is computed in an entirely similar manner as for LDPC codes. We nevertheless repeat the presentation of [7, Sec. II] for completeness and to introduce notation.

Consider first *systematic* RA codes. A variable node of degree d_v has d_v+1 incoming messages: d_v from the edge interleaver and one from the channel. The variable node decodes by adding L-values. The decoder outputs are

$$L_{i,\text{out}} = L_{\text{ch}} + \sum_{j \neq i} L_{j,\text{in}} \tag{1}$$

where $L_{j,\text{in}}$ is the jth *a priori* L-value going into the variable node, $L_{i,\text{out}}$ is the ith extrinsic L-value coming out of the variable node, and L_{ch} is the channel L-value.

Consider the AWGN channel with BPSK (± 1) modulation and noise variance σ_n^2. We define the normalized signal-to–noise ratio (SNR) as $E_b/N_0 = 1/(2R\sigma_n^2)$. This means that the variance of the channel L-value is (see the Appendix)

$$\sigma_{\text{ch}}^2 = \frac{4}{\sigma_n^2} = 8R \cdot \frac{E_b}{N_0}. \tag{2}$$

We model $L_{j,\text{in}}$ as the output L-value of an AWGN channel whose input is the jth interleaver bit transmitted using BPSK. The EXIT function of a degree–d_v variable node is then (see [7, Sec. II])

$$I_{E,\text{VND}}\left(I_A, d_v, \frac{E_b}{N_0}, R\right) = J\left(\sqrt{(d_v-1)\left[J^{-1}(I_A)\right]^2 + \sigma_{\text{ch}}^2}\right) \tag{3}$$

where the $J(\cdot)$ function is given in the Appendix.

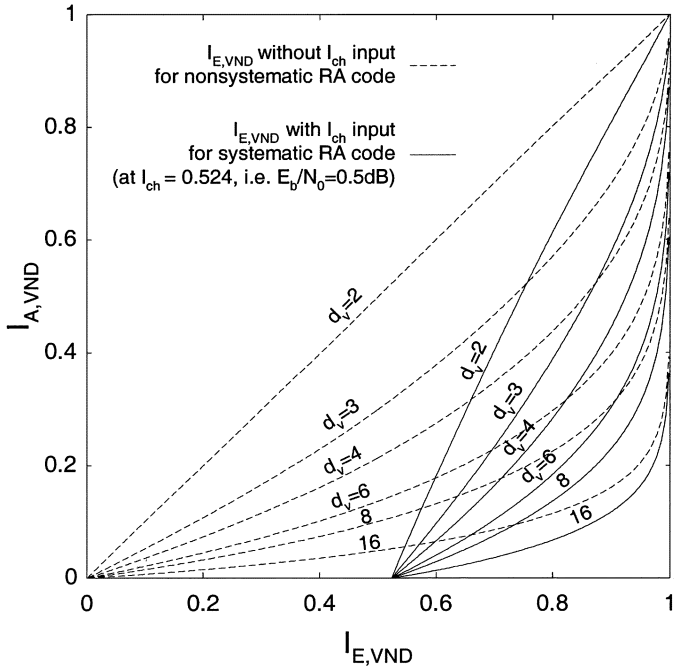


Fig. 3. VND EXIT curves for systematic RA codes (solid lines) and nonsystematic RA codes (dashed lines). The systematic curves are plotted assuming that $R = 1/2$.

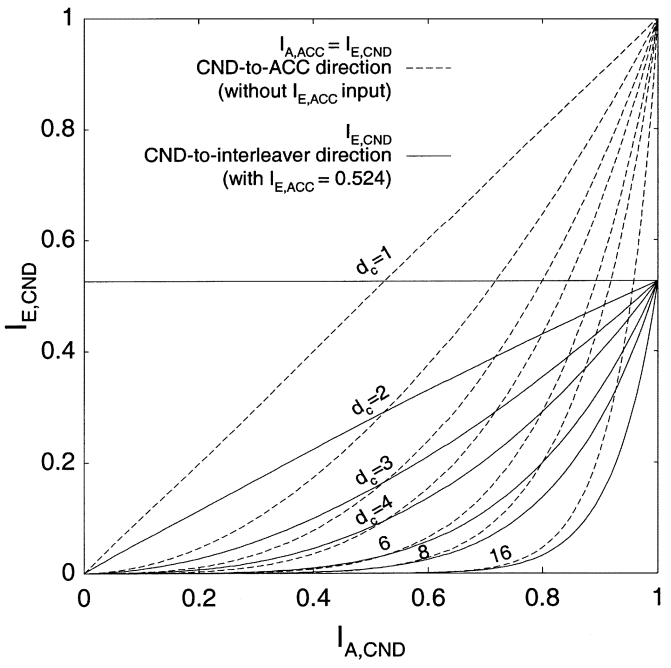


Fig. 4. CND EXIT curves for the CND-to-ACC direction (dashed lines) and the CND-to-interleaver direction (solid lines).

For *nonsystematic* RA codes, the variable node does not have a channel input. We must therefore set $\sigma_{\rm ch} = 0$ in (3) and obtain

$$I_{E,\rm VND}(I_A, d_v) = J\left(\sqrt{(d_v - 1)} \cdot J^{-1}(I_A)\right). \qquad (4)$$

Fig. 3 plots several variable node curves when $R = 1/2$ and $E_b/N_0 = 0.5$ dB (the figure has $I_{A,\rm VND} = I_A$). For *systematic* RA codes, the EXIT curves (solid lines) start at $I_{\rm ch} > 0$, where $I_{\rm ch}$ is the channel capacity at the $\sigma_{\rm ch}$ that is being considered. For *nonsystematic* RA codes, the EXIT curves (dashed lines) start at the origin. This means that if the ACC&CND curve also starts from the origin, the decoder will not begin converging. It is for this reason that we will use *check biregular* (or *doped*) RA codes.

D. EXIT Curve of the Inner Check Node Decoder

Consider next the CND. A check node of degree d_c has d_c+1 incoming messages: d_c from the edge interleaver and one from the accumulator (see Fig. 1). This differs from LDPC codes where all edges of the CND come from the edge interleaver. The decoding of a degree d_c check node is therefore the same as the decoding of a length $d_c + 1$ (or rate $d_c/(d_c + 1)$) single parity check code. The output L-values are (see [15, Sec.II.A] and [7, Sec. II])

$$L_{i,\rm out} = \ln \frac{1 - \prod_{j\neq i} \frac{1-e^{L_{j,\rm in}}}{1+e^{L_{j,\rm in}}}}{1 + \prod_{j\neq i} \frac{1-e^{L_{j,\rm in}}}{1+e^{L_{j,\rm in}}}} \qquad (5)$$

where the indexes i and j include the d_c messages from the edge interleaver and the message from the accumulator. In the L-value notation of [15], this is a "box-plus" operation:

$$L_{i,\rm out} = \sum_{j\neq i} \boxplus L_{j,\rm in}. \qquad (6)$$

As above, we model $L_{j,\rm in}$ as the output L-value of an AWGN channel whose input is the jth check node input bit transmitted using BPSK.

For further analysis, we consider separately the two directions of information flow through the CND: 1) The *CND-to-ACC direction*, where information is fed back from the interleaver to the ACC, and 2) the *CND-to-interleaver direction*, where information is fed forward through the edge interleaver to the VND.

CND-to-ACC Direction: We follow [7] (see also [16, p. 189] and [17]) and use a duality property for the binary erasure channel (BEC) that expresses the EXIT curve $I_{E,\rm SPC}(\cdot)$ of the length-d_c single parity check code in terms of the EXIT curve $I_{E,\rm REP}(\cdot)$ of the length d_c repeat code. The result is

$$I_{E,\rm SPC}(I_A, d_c) \approx 1 - I_{E,\rm REP}\left(1 - I_A, d_c\right). \qquad (7)$$

Simulations show this to be an accurate approximation; therefore, we write the EXIT curve of the CND-to-ACC direction as

$$\begin{aligned} I_{A,\rm ACC}(I_A, d_c) \approx& 1 - I_{E,\rm REP}\left(1 - I_A, d_c + 1\right) \\ =& 1 - J\left(\sqrt{d_c} \cdot J^{-1}\left(1 - I_A\right)\right). \end{aligned} \qquad (8)$$

CND-to-Interleaver Direction: Each CND node computes extrinsic information using its L-value from the ACC and $d_c - 1$ L-values from the edge interleaver. Using duality, we approximate the EXIT curve of the CND-to-interleaver direction as

$$\begin{aligned} & I_{E,\rm CND}\left(I_A, I_E, d_c\right) \\ \approx& 1 - I_{E,\rm REP}\left(1 - I_A, 1 - I_E, d_c\right) \\ =& 1 - J\left(\sqrt{(d_c - 1) \cdot \left[J^{-1}\left(1 - I_A\right)\right]^2 + \left[J^{-1}\left(1 - I_E\right)\right]^2}\right) \end{aligned} \qquad (9)$$

where $I_A = I_{A,\rm CND}$ and $I_E = I_{E,\rm ACC}$ (see Fig. 2).

Fig. 4 plots several check node EXIT curves. For the CND-to-interleaver direction, we have set $I_{E,\rm ACC} = 0.524$, which corresponds to $E_b/N_0 = 0.5$ dB and $R = 1/2$. Consider, e.g., the

case $d_c = 1$ so that there is one edge to the edge interleaver. From (8) and (9), we see that the CND-to-ACC direction has $I_{A,\mathrm{ACC}} = I_{A,\mathrm{CND}}$ and that the CND-to-interleaver direction has $I_{E,\mathrm{CND}} = I_{E,\mathrm{ACC}}$. Thus, the EXIT curves are a diagonal line and a horizontal line, respectively.

E. EXIT Curve of the Inner Accumulator Decoder

We compute the EXIT function of the accumulator decoder (a unit memory trellis decoder denoted ACC) by simulation and denote it by

$$I_{E,\mathrm{ACC}}\left(I_{A,\mathrm{ACC}}, \frac{E_b}{N_0}, R\right). \tag{10}$$

Note that this curve depends on E_b/N_0 and R. Some curves for $R = 1/2$ are plotted in Fig. 5. We point out that the purpose of the accumulator is to make $I_E(I_A = 1) = 1$, i.e., without the accumulator, the check node layer will have $I_E(I_A = 1) < 1$.

Fig. 5 also shows ACC curves for BECs that have the same capacity as the AWGN. These curves are given by (see [5, eq. (14)] or [17, eq. (27)])

$$I_{E,\mathrm{ACC}}(I_{A,\mathrm{ACC}}, q) = \left[\frac{1-q}{1-q\,I_{A,\mathrm{ACC}}}\right]^2 \tag{11}$$

where $1 - q = C(E_b/N_0)$, and $C(E_b/N_0)$ is the capacity in bits per use of the AWGN channel with binary antipodal modulation at the SNR E_b/N_0. For example, we have $C(E_b/N_0 = 0.5\ \mathrm{dB}) = 0.524$. We remark that the BEC curves differ sufficiently from the AWGN curves to warrant using the latter curves rather than (11).

F. EXIT Curve of the Combined Inner ACC and CND

We combine (8)–(10) to approximate the transfer curve $I_{E,\mathrm{ACC\,\&\,CND}}$ of the inner decoder comprising the ACC (a unit memory trellis decoder) and CND. The result is

$$\begin{aligned} &I_{E,\mathrm{ACC\,\&\,CND}}\left(I_{A,\mathrm{CND}}, d_c, \frac{E_b}{N_0}, R\right) \\ &= I_{E,\mathrm{CND}}\left(I_{A,\mathrm{CND}}, I_{E,\mathrm{ACC}}\left(I_{A,\mathrm{ACC}}, \frac{E_b}{N_0}, R\right), d_c\right) \\ &= I_{E,\mathrm{CND}}\left(I_{A,\mathrm{CND}},\right. \\ &\quad \left. I_{E,\mathrm{ACC}}\left(I_{A,\mathrm{ACC}}\left(I_{A,\mathrm{CND}}, d_c\right), \frac{E_b}{N_0}, R\right), d_c\right). \end{aligned} \tag{12}$$

Fig. 6 depicts several such EXIT curves. Note that the ACC curve is recovered by setting $d_c = 1$, i.e., $d_c = 1$ effectively turns off the check node decoder.

Observe that the curves with $d_c > 1$ in Fig. 6 and the nonsystematic VND curves in Fig. 3 start at the origin. This means that the decoder cannot even begin to converge, which led [5] to remark that nonsystematic RA codes are useless for $d_c > 1$. We remedy this defect by using a check node layer that is *biregular*, i.e., a nonzero fraction of the check nodes have degree one. This trick is akin to *code doping* [14], and it ensures that the decoder will make progress in the first iteration. Nonsystematic RA codes are thereby made practical and we show that they can in fact approach capacity.

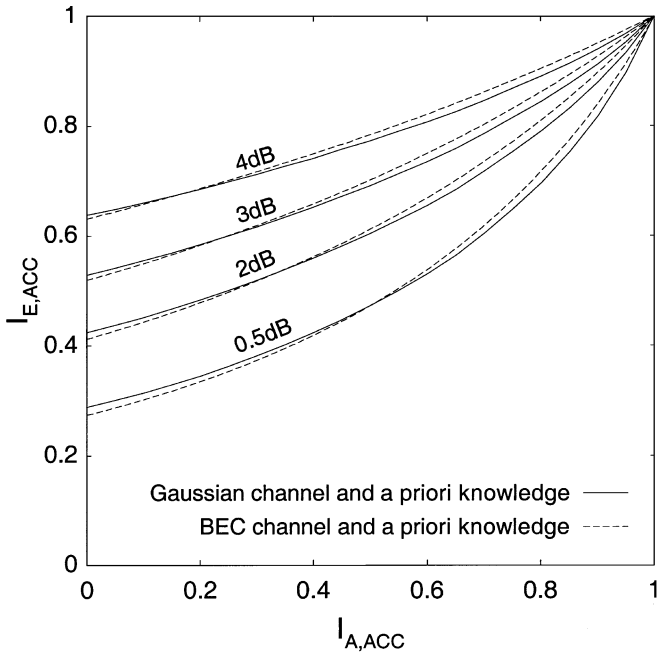


Fig. 5. EXIT curves for the ACC. A code rate of $R = 1/2$ was used to compute E_b/N_0.

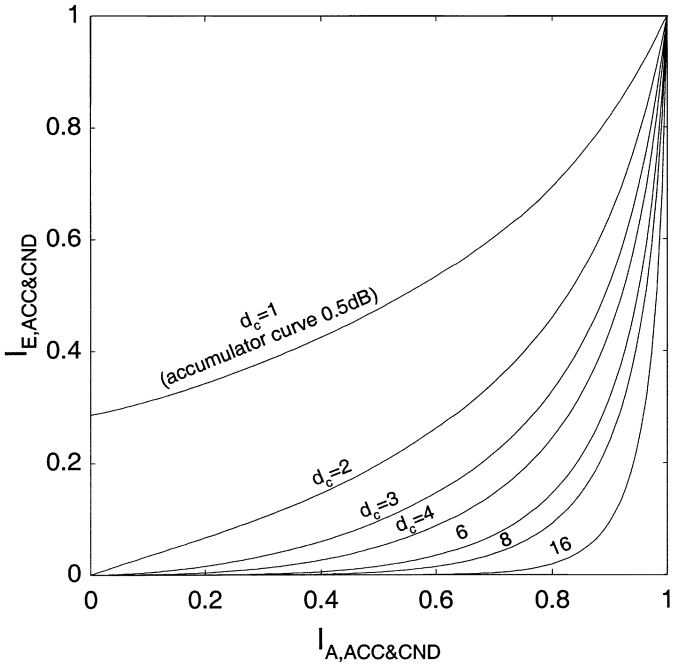


Fig. 6. EXIT curves for the combined ACC and CND. AWGN channel with $E_b/N_0 = 0.5$ dB at $R = 1/2$.

Finally, we remark that rather than measuring $I_{E,\mathrm{ACC}}$ and computing $I_{E,\mathrm{ACC\,\&\,CND}}$ via (12), one can alternatively measure $I_{E,\mathrm{ACC\,\&\,CND}}$ directly. This approach is practical because the latter curve will stay fixed for the curve fitting procedure described next.

G. EXIT Curves for Code Mixtures

We first consider *systematic* RA codes for which there are k variable nodes and $n - k$ check nodes. We will consider only

check-regular systematic RA codes, i.e., all check nodes have degree d_c. The remaining design involves specifying the variable node degrees $d_v^{(i)}$, $i = 1, \ldots, k$. Let D_v be the number of different variable node degrees, and denote these degrees by $\tilde{d}_{v,i}$, $i = 1, \ldots, D_v$. We follow [7], and write the average variable node degree as

$$\overline{d}_v = \sum_{i=1}^{D_v} a_{v,i} \cdot \tilde{d}_{v,i} \tag{13}$$

where $a_{v,i}$ is the fraction of *nodes* having degree $\tilde{d}_{v,i}$. Note that we have $\sum_i a_{v,i} = 1$. Since the number of interleaver edges at the VND and CND are the same, we further have $k\,\overline{d}_v = (n-k)\,d_c$ or

$$\overline{d}_v = \frac{1-R}{R} \cdot d_c. \tag{14}$$

Let $b_{v,i}$ be the fraction of *edges* incident to variable nodes of degree $\tilde{d}_{v,i}$. There are in total $(k\,a_{v,i})\,\tilde{d}_{v,i}$ edges involved with such nodes; therefore, we have

$$b_{v,i} = \frac{k\,a_{v,i}\,\tilde{d}_{v,i}}{k\,\overline{d}_v} = \frac{R\,\tilde{d}_{v,i}}{(1-R)\,d_c} \cdot a_{v,i}. \tag{15}$$

In [17] and [18], it is shown that the EXIT curve of a mixture of codes is an average of the component EXIT curves. As in [7], we must average using the $b_{v,i}$ (and not the $a_{v,i}$) because it is the *edges* that carry the extrinsic messages. The VND transfer curve is therefore given by

$$I_{E,\mathrm{VND}}\left(I_A, \frac{E_b}{N_0}, R\right) = \sum_{i=1}^{D_v} b_{v,i} \cdot I_{E,\mathrm{VND}}\left(I_A, \tilde{d}_{v,i}, \frac{E_b}{N_0}, R\right). \tag{16}$$

Note that only $D_v - 2$ edge fractions can be adjusted because we must enforce (14) and $\sum_i b_{v,i} = 1$. Thus, we must choose $D_v \geq 3$ to permit curve fitting. We will see that $D_v = 3$ already gives good results.

We now turn to *nonsystematic* RA codes for which there are n check nodes (rather than $n-k$ check nodes as for systematic RA codes). As already stated above, we overcome the initial convergence problems by using a CND layer that is *biregular*. We write this as $D_c = 2$ and compute the average check node degree as

$$\overline{d}_c = \sum_{i=1}^{D_c} a_{c,i} \cdot \tilde{d}_{c,i} = a_{c,1} + a_{c,2} \cdot \tilde{d}_{c,2} \tag{17}$$

where $a_{c,i}$ is the fraction of check nodes with degree $\tilde{d}_{c,i}$, and where we have set $\tilde{d}_{c,1} = 1$. The number of interleaver edges is $k\,\overline{d}_v = n\,\overline{d}_c$ so that

$$\overline{d}_v = \frac{\overline{d}_c}{R}. \tag{18}$$

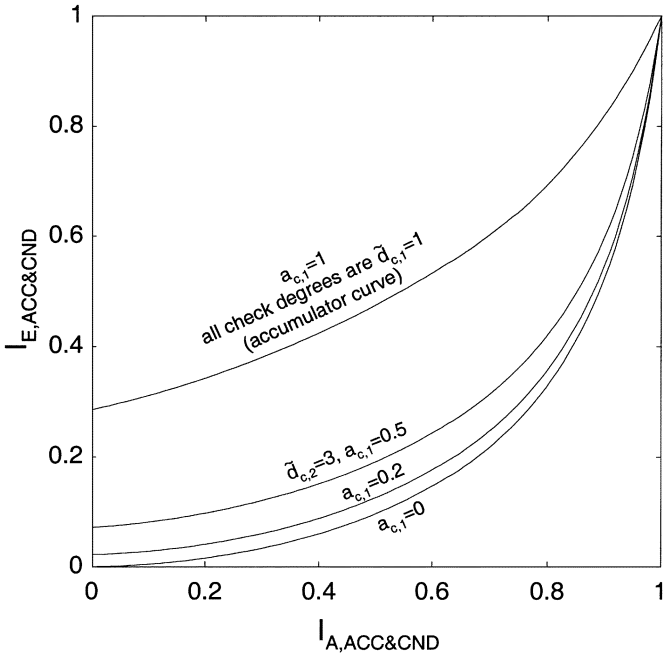


Fig. 7. EXIT curves for an ACC and a biregular CND. The channel parameters are $E_b/N_0 = 0.5$ dB and $R = 1/2$.

The fraction $b_{v,i}$ of edges incident to variable nodes of degree i is now

$$b_{v,i} = \frac{R\,\tilde{d}_{v,i}}{\overline{d}_c} \cdot a_{v,i}. \tag{19}$$

The EXIT curves of the biregular CND can now be obtained either by Monte Carlo simulation or by using

$$I_{E,\mathrm{ACC\,\&\,CND}}\left(I_A, \frac{E_b}{N_0}, R\right) = \sum_{i=1}^{2} b_{c,i} \cdot I_{E,\mathrm{ACC\,\&\,CND}}\left(I_A, \tilde{d}_{c,i}, \frac{E_b}{N_0}, R\right) \tag{20}$$

where $b_{c,i} = (\tilde{d}_{c,i} \,/\, \overline{d}_c) \cdot a_{c,i}$. We use (20) with (8)–(12) so that we need to simulate only the EXIT curve of the ACC.

Fig. 7 plots several EXIT curves for a biregular CND layer. Observe that increasing $a_{c,1}$ lifts the EXIT curve to open a convergence tunnel (cf. [14]). Observe also that for $a_{c,1} = 1$, we recover the ACC EXIT curve (see Fig. 5).

H. Design Examples

We proceed as in [7] and fit the EXIT curve of the VND to that of the ACC & CND. We consider $R = 1/2$ and BPSK on AWGN channels. The E_b/N_0 capacity limit for this setup is at 0.19 dB. We choose $D_v = 3$ and perform the curve fitting at 0.5 dB. Choosing $D_v > 3$ would let us approach capacity even more closely.

Fig. 8 shows a curve fit for a *systematic* RA code with $d_c = 6$ and

$$\begin{aligned} \tilde{d}_{v,1} &= 2, \quad a_{v,1} = 0.0631, \quad b_{v,1} = 0.0210 \\ \tilde{d}_{v,2} &= 3, \quad a_{v,2} = 0.6306, \quad b_{v,2} = 0.3153 \\ \tilde{d}_{v,3} &= 13, \; a_{v,3} = 0.3063, \quad b_{v,3} = 0.6637. \end{aligned}$$

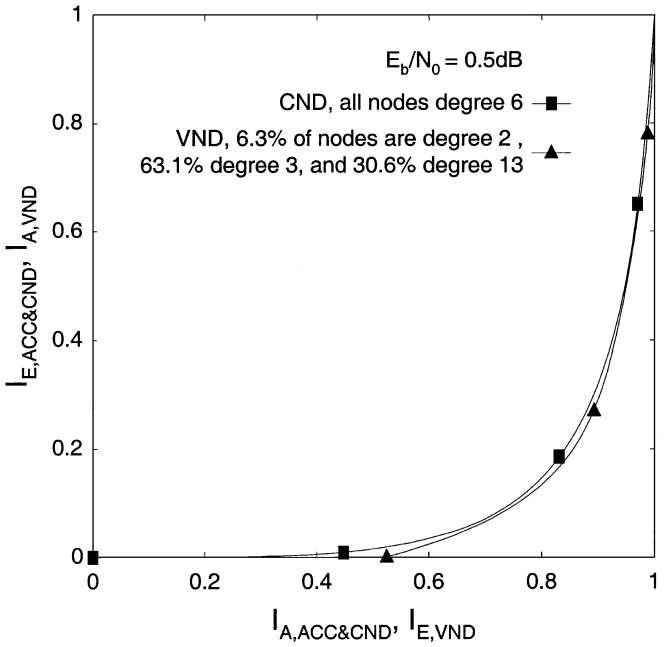


Fig. 8. Curve fit for a systematic RA code with $R = 1/2$.

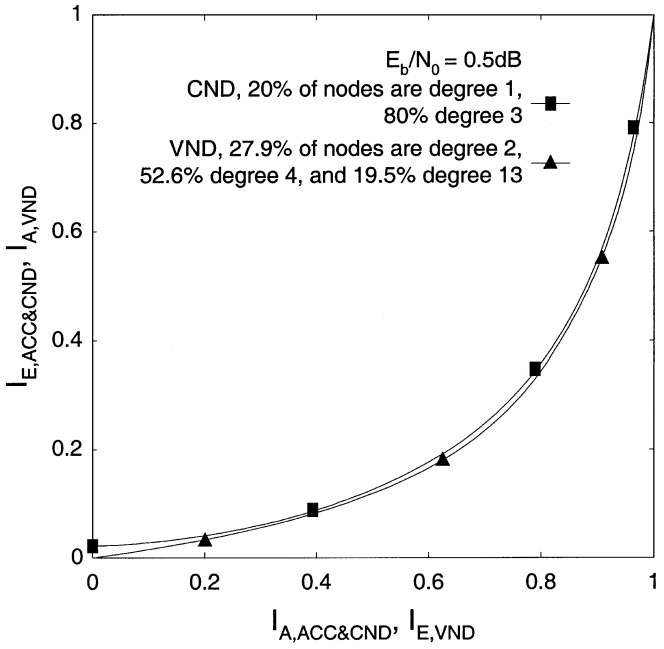


Fig. 9. Curve fit for a nonsystematic RA code with $R = 1/2$.

Fig. 9 shows a curve fit for a *nonsystematic* RA code with a *biregular* CND layer. The code's parameters are $a_{c,1} = 0.2$, $\tilde{d}_{c,1} = 1$, $\tilde{d}_{c,2} = 3$, and

$$\begin{aligned} \tilde{d}_{v,1} &= 2, \quad a_{v,1} = 0.2789, \quad b_{v,1} = 0.1073 \\ \tilde{d}_{v,2} &= 4, \quad a_{v,2} = 0.5258, \quad b_{v,2} = 0.4044 \\ \tilde{d}_{v,3} &= 13, \quad a_{v,3} = 0.1953, \quad b_{v,3} = 0.4883. \end{aligned}$$

Both the systematic and nonsystematic RA codes have convergence thresholds at about 0.5 dB. Furthermore, simulations with length $n = 10^5$ codes and 100 iterations show a turbo cliff at about 0.5 dB (we measure the turbo cliff at a bit error rate (BER) of 10^{-4}).

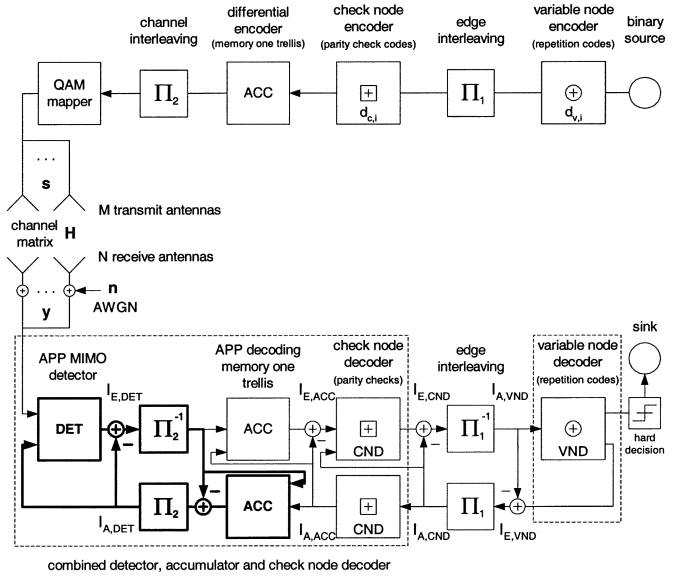


Fig. 10. MIMO model.

The usual error "floor" was observed near a BER of 10^{-5} when many blocks were simulated. This floor occurs for several reasons: the large percentage of degree–2 variable nodes makes the code's distance spectrum poor and makes the EXIT convergence tunnel narrow near $(I_A, I_E) = (1, 1)$, the limited girth of our interleavers introduces correlations between the extrinsic L-values, and the BEC stopping sets [11], [19] cause the convergence to slow for the AWGN channel as well. We chose our interleavers to avoid BEC stopping sets of size one, i.e., we ensured that there were no parallel edges between nodes. We remark that the error floor can be lowered by increasing the degrees of the variable and check nodes, but this also increases the number of edges and, hence, the decoding complexity.

Our simulations show that check-biregular, nonsystematic RA codes can perform as well as check-regular, systematic RA codes. Moreover, the nonsystematic RA codes are easier to design for iterative *detection and decoding*, as discussed next.

III. RA Codes and MIMO Communication

Systematic RA decoders receive channel L-values at both the inner and outer decoder. This complicates the design of an iterative detection and decoding scheme because one must take the detector into account for both the inner and outer EXIT curves. In contrast, for *nonsystematic* RA codes, only the inner EXIT curve depends on the detector. Furthermore, we now know that nonsystematic RA codes can perform as well as systematic RA codes. We will therefore combine MIMO detectors with the inner decoder (comprising the ACC and a check-biregular CND layer) of a nonsystematic RA code to form a new inner decoder. We then design the VND transfer curve to match the compound inner detector/ACC/CND transfer curve.

A. MIMO Communication Link

Consider the multiantenna setup of [7, Sec. III] that is depicted in Fig. 10. We again repeat the description of [7] for completeness and to introduce notation.

Suppose there are M transmit and N receive antennas. Each transmitter symbol is thus an $M\times 1$ vector $\mathbf{s}=[s_1,\ldots,s_M]^T$ whose entries take on complex values in a constellation set. We consider constellations of size 2^{M_c} so that each symbol carries $M\cdot M_c$ coded bits. For example, for quadrature phase-shift keying (QPSK), we have $M_c=2$. The average energy per transmit symbol is limited to E_s, and we assume that $E[\|s_m\|^2]=E_s/M$.

The receiver sees $N\times 1$ vectors $\mathbf{y}=\mathbf{Hs}+\mathbf{n}$, where $\mathbf{H}$ is the $N\times M$ channel matrix, and $\mathbf{n}$ is a $N\times 1$ noise vector. We assume the entries of $\mathbf{n}$ to be independent, complex, zero-mean, Gaussian random variables with independent real and imaginary parts each having variance $\sigma^2=N_0/2$. We define the normalized SNR E_b/N_0 to be

$$\left.\frac{E_b}{N_0}\right|_{\rm dB}=\left.\frac{E_s}{N_0}\right|_{\rm dB}+10\log_{10}\frac{N}{RMM_c}. \tag{21}$$

We assume $\mathbf{H}$ to be known by the receiver only and consider a Rayleigh fading channel so that the entries of $\mathbf{H}$ are independent, complex, zero-mean, Gaussian random variables with independent real and imaginary parts, each having variance 1/2 [26, Sec. 4]. For a *quasistatic* channel, the matrix $\mathbf{H}$ remains unchanged over long time intervals, whereas for an *ergodic* channel, $\mathbf{H}$ changes for every symbol $\mathbf{s}$. We consider only the ergodic model whose capacity is (see [25]–[27])

$$C=E\left[\log_2\det\left(\mathbf{I}+\frac{E_s}{N_0}\frac{1}{M}\mathbf{H}\mathbf{H}^\dagger\right)\right] \tag{22}$$

where $\mathbf{I}$ is the identity matrix, and $\mathbf{H}^\dagger$ denotes the complex-conjugate transpose of $\mathbf{H}$. One achieves capacity by using Gaussian distributed symbols $\mathbf{s}$. However, we consider only Gray-mapped QPSK.

B. EXIT Curve of the MIMO Detector

The MIMO detector performs APP bit detection by considering all 2^{MM_c} possible hypotheses on $\mathbf{s}$ (see [20]–[24]). The detector EXIT curve cannot be described in closed form; therefore, we measure it by Monte Carlo simulation [7], [28]. We denote the detector by DET and its EXIT curve by

$$I_{E,\rm DET}\left(I_{A,\rm DET},\frac{E_b}{N_0}\right). \tag{23}$$

The detector's soft output is forwarded to the decoder, which in turn computes extrinsic values to be used as *a priori* knowledge by the detector for the next iteration.

C. EXIT Curve of the Combined Detector, ACC, and CND

We connect the detector to the inner RA decoder comprising the ACC and CND via an interleaver, as shown in Fig. 10. The detector takes channel vectors $\mathbf{y}$ and computes *a posteriori* L-values that are deinterleaved and fed to the ACC. The purpose of the interleaver is the same as the RA code edge interleaver, namely, to break local L-value dependencies.

The receiver chain of Fig. 10 is in fact a double serial concatenation (see [29]). Such a structure has the peculiar feature

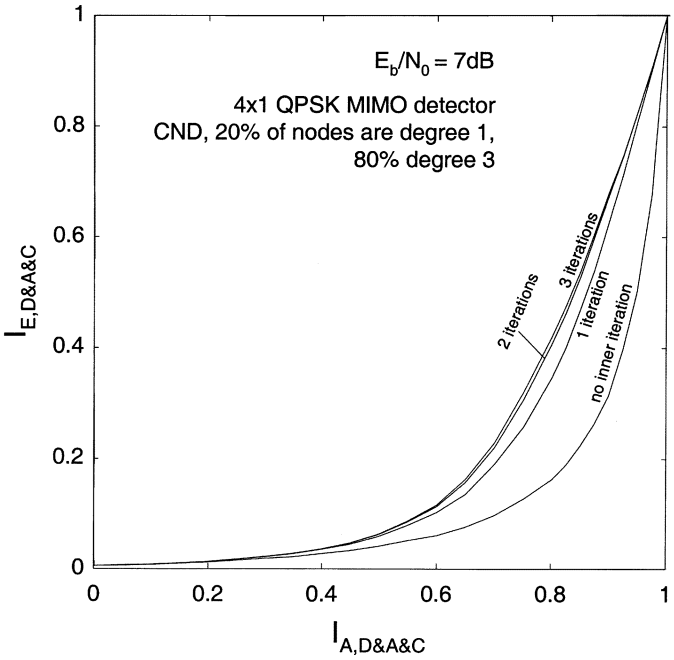


Fig. 11. EXIT curves for an inner 4×1–MIMO detector, ACC and CND after several inner DET/ACC iterations. Parameters: $R=1/2$, $E_b/N_0=7$ dB, $n=10^5$, $a_{c,1}=0.2$, $\bar{d}_{c,1}=1$, and $\bar{d}_{c,2}=3$.

of requiring an *inner detection loop*, i.e., iterations between the detector and the ACC (shown as thick lines in Fig. 10). The purpose of this loop is to adequately infuse the *a priori* knowledge $I_{A,\rm ACC}$. We found that at most three inner iterations suffice to make $I_{E,\rm ACC}$ close to its maximum given the available $I_{A,\rm ACC}$. For example, Fig. 11 shows the combined DET, ACC, and CND curves after different numbers of inner iterations for a 4×1–MIMO detector. Observe that we achieve saturation after three iterations; other scenarios, like a 2×1– or 4×2–MIMO detector, require only two iterations to saturate. As in this figure, we denote the overall transfer curve of the inner decoder as $I_{E,D\&A\&C}(I_{A,D\&A\&C})$, or, simply, as $I_{E,\rm CND}(I_{A,\rm CND})$.

An alternative to the structure of Fig. 10 is to omit the interleaver Π_2 and to perform the detection and accumulator decoding in a single trellis (DET & ACC trellis). The result is a new code for which the trellis has two states and 2^{MM_c} incoming/outgoing edges per state, each edge being labeled by the MM_c bits corresponding to one vector-channel symbol. The transition from one state to the next is determined by running MM_c bits through the differential encoder, with 2^{MM_c-1} transitions arriving in state 0 and 2^{MM_c-1} transitions arriving in state 1. For detection/decoding, one can apply the Bahl, Cocke, Jelinek, Raviv (BCJR) algorithm [30], [31] to the trellis.

The code design for this system is the same as before, i.e., we simulate to determine the DET & ACC curve, and then fit the VND curve to the DET & ACC curve. However, the *trellis detection* has two advantages over the system of Fig. 10. First, no inner iterations are required, which reduces decoding complexity; observe that the detector must perform the same APP bit detection for either scheme. Second, simulations show that trellis detection is more robust against code parameter mismatches. For instance, suppose an RA code is designed for the AWGN channel but is used over a QPSK 4×1–MIMO channel;

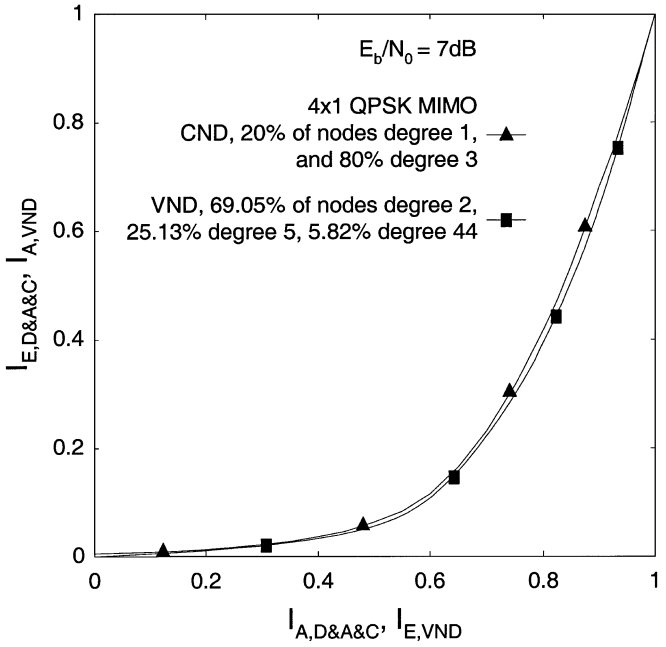


Fig. 12. Curve fit for a QPSK 4×1–MIMO detector and a nonsystematic, check-biregular RA code with $R = 1/2$.

then, the performance penalty owing to a VND-curve mismatch is smaller with a trellis detector than with the system of Fig. 10. This is discussed in more detail in [32].

These advantages suggest that one should use MIMO trellis detection rather than an inner detection loop. However, for some applications, one might have only a separate detector and interleaver available, e.g., on an application specific integrated circuit (ASIC), which does not allow for joint trellis processing. We have thus chosen to illustrate our approach with the more modular structure of Fig. 10. The next section considers codes designed for a double serial concatenation.

D. Design Examples for Combined Detection and Decoding

We choose $a_{c,1} = 0.2$, $\tilde{d}_{c,1} = 1$, and $\tilde{d}_{c,2} = 3$. The inner decoder/detector curve $I_{E,D\&A\&C}(I_{A,D\&A\&C})$ then depends only on the channel and not on the remaining code design. We set $R = 1/2$ and again restrict ourselves to just three different variable node degrees ($D_v = 3$). Fig. 12 plots a resulting EXIT chart for a 4×1 MIMO channel.

Table I shows our simulation results. The code length for the simulations was $n = 10^5$, and the decoder performed 100 iterations (due to the complexity of the APP processing the number of simulated blocks for the $4 \times N$ cases was limited to 40). Comparing the results with [7], we find that our nonsystematic RA codes perform as well as LDPC codes. The error floor behavior of both code families was similar. We have also listed the performance of a Universal Mobile Telecommunications System (UMTS) standard turbo code as parallel concatenated code (PCC). The PCC has length $n = 10^5$, and memory 3 constituent codes with feedback and feedforward generator polynomials $1+D^2+D^3$ and $1+D+D^3$, respectively. Observe that RA codes outperform this code by a wide margin when there are more transmit than receive antennas. This effect is explained in [7] and [28].

TABLE I
PARAMETERS FOR $M \times N$ MIMO USING QPSK WITH $R = 1/2$ RA CODES. OVERALL RATE: M BITS/CHANNEL USE.

2×1		
Capacity E_b/N_0 at 3.25dB		
Curve fit at 3.50dB		
BER 10^{-4} at 3.95dB		
(PCC at 5.00dB)		
$\tilde{d}_{v,i}$	a_i [%]	b_i [%]
2	37.66	14.48
4	50.13	38.57
20	12.21	46.95

4×1			4×2		
Capacity E_b/N_0 at 6.65dB			Capacity E_b/N_0 at 2.95dB		
Curve fit at 7.00dB			Curve fit at 3.30dB		
BER 10^{-4} at 7.90dB			BER 10^{-4} at 3.60dB		
(PCC at 13.8dB)			(PCC at 5.40dB)		
$\tilde{d}_{v,i}$	a_i [%]	b_i [%]	$\tilde{d}_{v,i}$	a_i [%]	b_i [%]
2	69.05	26.56	2	41.84	16.09
5	25.13	24.16	4	48.46	37.28
44	5.82	49.28	25	9.70	46.63

4×3			4×4		
Capacity E_b/N_0 at 1.97dB			Capacity E_b/N_0 at 1.47dB		
Curve fit at 2.30dB			Curve fit at 1.80dB		
BER 10^{-4} at 2.40dB			BER 10^{-4} at 1.90dB		
(PCC at 3.20dB)			(PCC at 2.40dB)		
$\tilde{d}_{v,i}$	a_i [%]	b_i [%]	$\tilde{d}_{v,i}$	a_i [%]	b_i [%]
2	36.75	14.13	2	35.39	13.61
4	51.87	39.90	4	50.99	39.22
21	11.38	45.97	18	13.62	47.17

IV. CONCLUDING REMARKS

We introduced a new class of RA codes that are nonsystematic, check–biregular, and that can approach capacity on AWGN and MIMO fading channels. We showed how to design degree distributions for these as well as the usual systematic RA codes. The design was based on EXIT charts, and simulations verified that all our codes operate within about 1 dB of their respective capacity limits at a BER of 10^{-4}. Note that other design methods for RA codes were recently presented in [33].

There are several promising avenues for further exploration of RA codes. For example, the degree distributions listed in

Table I can certainly be improved upon. Other open problems are finding degree distributions that achieve low error floors with specific modulations and understanding that graphical decoding structures are robust to channel variations.

APPENDIX

Consider $Y = X + N$, where $\Pr(X = \pm 1) = 1/2$, and N is zero-mean, Gaussian noise with variance σ_n^2. The channel L-value is

$$L_{\rm ch}(y) = \log \frac{p(y|x=+1)}{p(y|x=-1)} = \frac{2}{\sigma_n^2}\, y. \tag{24}$$

The mean and variance of $L_{\rm ch}(Y)$ conditioned on $X = \pm 1$ are the respective

$$\mu_{\rm ch} = \frac{\pm 2}{\sigma_n^2} = \pm 4R \cdot \frac{E_b}{N_0} \tag{25}$$

$$\sigma_{\rm ch}^2 = \frac{4}{\sigma_n^2} = 8R \cdot \frac{E_b}{N_0}. \tag{26}$$

The $J(\cdot)$–function gives the mutual information between X and $L_{\rm ch}(Y)$. We compute

$$\begin{aligned} J(\sigma_{\rm ch}) &= H(X) - H(X|L_{\rm ch}(Y)) \\ &= 1 - \int_{-\infty}^{\infty} \frac{e^{-\left(\xi - \sigma_{\rm ch}^2/2\right)^2/2\sigma_{\rm ch}^2}}{\sqrt{2\pi\sigma_{\rm ch}^2}} \cdot \log_2\left[1 + e^{-\xi}\right] d\xi \end{aligned} \tag{27}$$

where $H(X)$ is the entropy of X, and $H(X|L_{\rm ch}(Y))$ is the entropy of X conditioned on $L_{\rm ch}(Y)$. We can use $J(\cdot)$ to compute the capacity of BPSK on an AWGN channel: Note that $I(X; L_{\rm ch}(Y))$ is the same as $I(X;Y)$. The capacity is therefore $J(\sigma_{\rm ch}) = J\left(2/\sigma_n\right)$. For computer implementation, we used the same approximation to the $J(\cdot)$–function as in [7].

REFERENCES

[1] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. 36th Annu. Allerton Conf. Commun., Control, Computing*, Sept. 1998, pp. 201–210.

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

[3] R. G. Gallager, "Low-density parity-check codes," Sc.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1960.

[4] ——, "Low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. IT-8, pp. 21–28, Jan. 1962.

[5] H. Jin, A. Khandekar, and R. McEliece, "Irregular repeat-accumulate codes," in *Proc. 2nd Int. Symp. Turbo Codes Related Topics*, Brest, France, Sept. 2000, pp. 1–8.

[6] S. ten Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.

[7] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, submitted for publication.

[8] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 569–584, Feb. 2001.

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

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

[11] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 638–656, Feb. 2001.

[12] B. J. Frey and D. J. C. MacKay, "Irregular turbo codes," in *Proc. 37th Ann. Allerton Conf. Commun., Control, Computing*, Oct. 1999.

[13] J. Boutros, G. Caire, E. Viterbo, H. Saway, and S. Vialle, "Turbo code at 0.03 dB from capacity limit," *Proc. IEEE Int. Symp. Inform. Theory*, p. 56, July 2002.

[14] S. ten Brink, "Code doping for triggering iterative decoding convergence," *Proc. IEEE Int. Symp. Inform. Theory*, p. 235, July 2001.

[15] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 429–445, Mar. 1996.

[16] S. Y. Chung, "On the Construction of Some Capacity-Approaching Coding Schemes," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2000.

[17] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: A model and two properties," in *Proc. 36th Ann. Conf. Informaiton Sciences Syst.*, Princeton, Mar. 2002.

[18] M. Tuechler and J. Hagenauer, "EXIT charts and irregular codes," in *Proc. 36th Annu. Conf. Inform. Sciences Syst.*, Princeton, NJ, Mar. 2002.

[19] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1570–1579, June 2002.

[20] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. Veh. Technol. Conf.*, Sept. 2000, pp. 473–478.

[21] A. van Zelst, R. van Nee, and G. Awater, "Turbo-BLAST and its performance," in *Proc. Veh. Technol. Conf.*, May 2001.

[22] A. Stefanov and T. Duman, "Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: System model, decoding approaches, and practical considerations," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 958–968, May 2001.

[23] C. Schlegel and A. Grant, "Concatenated space-time coding," in *Proc. PIMRC*, San Diego, CA, Sept. 2001.

[24] B. M. Hochwald and S. ten Brink, "Iterative list sphere decoding to attain capacity on a multi-antenna link," in *39th Annu. Allerton Conf. Commun., Contr., Comput.*, Oct. 2001.

[25] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[26] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[27] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[28] S. ten Brink and B. M. Hochwald, "Detection thresholds of iterative MIMO processing," *Proc. IEEE Int. Symp. Inform. Theory*, p. 22, July 2002.

[29] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Analysis, design, and iterative decoding of double serially concatenated codes with interleavers," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 231–244, Feb. 1998.

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

[31] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. Int. Contr. Conf.*, June 1995, pp. 1009–1013.

[32] S. ten Brink and G. Kramer, "Turbo processing for scalar and vector channels," in *Proc. 3rd Int. Symp. Turbo Codes Related Topics*, Brest, France, Sept. 2003, pp. 23–30.

[33] A. Roumy, S. Guemghar, G. Caire, and S. Verdú, "Design methods for irregular repeat-accumulate codes," *Proc. IEEE Int. Symp. Inform. Theory*, p. 2, June-July.

Stephan ten Brink received the Dipl.-Ing. degree in electrical engineering from the University of Stuttgart, Stuttgart, Germany, in 1997. From 1997 to 2000, he was a research assistant with the Institute of Telecommunications, Stuttgart, Germany, where he was pursuing the doctoral degree.

Since November 2000, he has been a Member of Technical Staff with the Wireless Research Lab., Bell Laboratories, Lucent Technologies, Holmdel, NJ. His research interests include multiuser detection, error-correcting coding, and channel estimation for digital communication systems.

Gerhard Kramer (S'91–M'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Manitoba, Winnipeg, MB, Canada, in 1991 and 1992, respectively, and the Dr. Sc. Techn. degree from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 1998.

From August 1998 to March 2000, he was with Endora Tech AG, Basel, Switzerland, as a communications engineering consultant. Since May 2000, he has been a Member of Technical Staff with the Mathematics of Communications Research Department, Bell Laboratories, Lucent Technologies, Murray Hill, NJ.