Coding Theorems for "Turbo-Like" Codes*

Dariush Divsalar, Hui Jin, and Robert J. McEliece
Jet Propulsion Laboratory and California Institute of Technology
Pasadena, California USA
E-mail: dariush@shannon.jpl.nasa.gov, (hui, rjm)@systems.caltech.edu

Abstract.

In this paper we discuss AWGN coding theorems for ensembles of coding systems which are built from fixed convolutional codes interconnected with random interleavers. We call these systems "turbo-like" codes and they include as special cases both the classical turbo codes [1,2,3] and the serial concatentation of interleaved convolutional codes [4]. We offer a general conjecture about the behavior of the ensemble (maximum-likelihood decoder) word error probability as the word length approches infinity. We prove this conjecture for a simple class of rate $1/q$ *serially concatenated codes where the outer code is a q-fold repetition code and the inner code is a rate* 1 *convolutional code with transfer function* $1/(1+D)$. *We believe this represents the first rigorous proof of a coding theorem for turbo-like codes.*

1. Introduction.

The 1993 discovery of turbo codes by Berrou, Glavieux, and Thitimajshima [1] has revolutionized the field of error-correcting codes. In brief, turbo codes have enough randomness to achieve reliable communication at data rates near capacity, yet enough structure to allow practical encoding and decoding algorithms. This paper is an attempt to illuminate the first of these two attributes, i.e., the "near Shannon limit" capabilities of turbo-like codes on the AWGN channel.

Our specific goal is to prove AWGN coding theorems for a class of generalized concatenated convolutional coding systems with interleavers, which we call "turbo-like" codes. This class includes both parallel concatenated convolutional codes (classical turbo codes) [1, 2, 3] and serial concatenated convolutional codes [4] as special cases. Beginning with a code structure of this type, with fixed component codes and interconnection topology, we attempt to show that as the block length approaches infinity, the ensemble (over all possible interleavers) maximum likelihood error probability approaches zero if E_b/N_0 exceeds some threshold. Our proof technique is to derive an explicit expression for the ensemble input-output weight enumerator (IOWE) and then to use this expression, in combination with either the classical union bound, or the recent "improved" union bound of Viterbi and Viterbi [9], to show that the maximum likelihood word error probability approaches zero as $N \to \infty$. Unfortunately the difficulty of the first step, i.e., the computation of the ensemble IOWE, has kept us from full success, except for some very simple coding systems, which we call *repeat and accumulate* codes. Still, we are optimistic that this technique will yield coding theorems for a much wider class of interleaved concatenated codes. In any case, it is satisfying to have rigorously proved coding theorems for even a restricted class of turbo-like codes.

Here is an outline of the paper. In Section 2 we quickly review the classical union bound on maximum-likelihood word error probability for block codes on the AWGN

* Dariush Divsalar's work, and a portion of Robert McEliece's work, was performed at JPL under contract with NASA. The remainder of McEliece's work, and Hui Jin's work, was performed at Caltech and supported by NSF grant no. NCR-9505975, AFOSR grant no. 5F49620-97-1-0313, and grant from Qualcomm.

channel, which is seen to depend on the code's weight enumerator. In Section 3 we define the class of "turbo-like" codes, and give a formula for the average input-output weight enumerator for such a code. In Section 4 we state a conjecture (the interleaver gain exponent conjecture) about the ML decoder performance of turbo-like codes. In Section 5, we define a special class of turbo-like codes, the repeat-and-accumulate codes, and prove the IGE conjecture for them. Finally, in Section 6 we present performance curves for some RA codes, using an iterative, turbo-like, decoding algorithm. This performance is seen to be remarkably good, despite the simplicity of the codes and the suboptimality of the decoding algorithm.

2. Union Bounds on the Performance of Block Codes.

In this section we will review the classical union bound on the maximum-likelihood word error probability for block codes.

Consider a binary linear (n, k) block code C with code rate $r = k/n$. The *(output) weight enumerator* (WE) for C is the sequence of numbers $A_0, \ldots, A_n$, where A_h denotes the number of codewords in C with (output) weight h. The *input-output weight enumerator* (IOWE) for C is the array of numbers $A_{w,h}$, $w = 0, 1, \ldots, k$, $h = 0, 1, \ldots, n$: $A_{w,h}$ denotes the number of codewords in C with input weight w and output weight h.

The union bound on the word error probability P_W of the code C over a memoryless binary-input channel, using maximum likelihood decoding, has the well-known form

$$P_W \leq \sum_{h=1}^{n} A_h z^h \tag{2.1}$$

$$= \sum_{h=1}^{n} \left(\sum_{w=1}^{k} A_{w,h} \right) z^h. \tag{2.2}$$

In (2.1) and (2.2), the function z^h represents an upper bound on the pairwise error probability for two codewords separated by Hamming (output) distance h. For AWGN channels, $z = e^{-rE_b/N_0}$ where E_b/N_0 is the signal-to-noise ratio per bit.

3. The Class of "Turbo-Like" Codes.

In this section, we consider a general class of concatenated coding systems of the type depicted in Figure 1, with q encoders (circles) and $q-1$ interleavers (boxes). The ith code C_i is an (n_i, N_i) linear block code, and the ith encoder is preceded by an interleaver (permuter) P_i of size N_i, except C_1 which is not preceded by an interleaver, but rather is connected to the input. The overall structure must have no loops, i.e., it must be a graph-theoretic tree. We call a code of this type a "turbo-like" code.

Define $s_q = \{1, 2, \ldots, q\}$ and subsets of s_q by $s_I = \{i \in s_q : C_i \text{ connected to input}\}$, $s_O = \{i \in s_q : C_i \text{ connected to output }\}$, and its complement $\overline{s}_O$. The overall system depicted in Figure 1 is then an encoder for an (n, N) block code with $n = \sum_{i \in s_O} n_i$.

If we know the IOWE $A^{(i)}_{w_i,h_i}$'s for the constituent codes C_i, we can calculate the *average* IOWE $A_{w,h}$ for the overall system (averaged over the set of all possible interleavers), using the *uniform interleaver* technique [2]. (A uniform interleaver is defined as a probabilistic device that maps a given input word of weight w into all distinct $\binom{N_i}{w}$ permutations of it with equal probability $p = 1/\binom{N_i}{w}$.) The result is

$$A_{w,h} = \sum_{\substack{h_i : i \in s_O \\ \Sigma h_i = h}} \sum_{h_i : i \in \overline{s}_O} A^{(1)}_{w_1,h_1} \prod_{i=2}^{q} \frac{A^{(i)}_{w_i,h_i}}{\binom{N_i}{w_i}} \tag{3.1}$$

In (3.1) we have $w_i = w$ if $i \in s_I$, and $w_i = h_j$ if C_i is preceeded by C_j (see Figure 2.). We do not give a proof of formula (3.1), but it is intuitively plausible if we note that the term $A^{(i)}_{w_i,h_i}/\binom{N_i}{w_i}$ is the probability that a random input word to C_i of weight w_i will produce an output word of weight h_i.

For example, for the $(n_2+n_3+n_4, N)$ encoder of Figure 1 the formula (3.1) becomes

$$
\begin{aligned}
A_{w,h} &= \sum_{\substack{h_1,h_2,h_3,h_4 \\ (h_2+h_3+h_4=h)}} A^{(1)}_{w_1,h_1} \frac{A^{(2)}_{w_2,h_2}}{\binom{N_2}{w_2}} \frac{A^{(3)}_{w_3,h_3}}{\binom{N_3}{w_3}} \frac{A^{(4)}_{w_4,h_4}}{\binom{N_4}{w_4}} \\
&= \sum_{\substack{h_1,h_2,h_3,h_4 \\ (h_2+h_3+h_4=h)}} A^{(1)}_{w,h_1} \frac{A^{(2)}_{w,h_2}}{\binom{N}{w}} \frac{A^{(3)}_{h_1,h_3}}{\binom{n_1}{h_1}} \frac{A^{(4)}_{h_1,h_4}}{\binom{n_1}{h_1}}.
\end{aligned}
$$

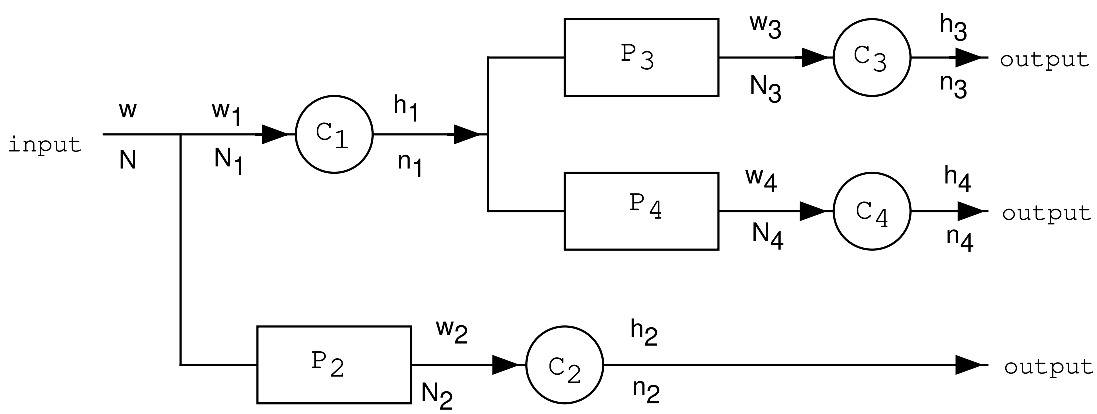

Figure 1. A "turbo-like" code with $s_I = \{1,2\}$, $s_O = \{2,3,4\}$, $\bar{s}_O = \{1\}$.

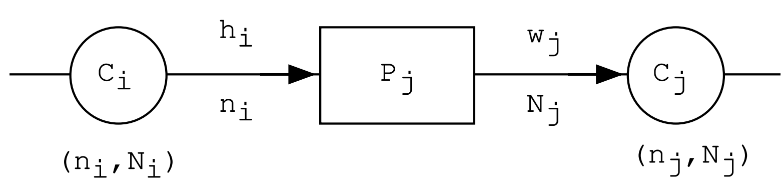

Figure 2. C_i (an (n_i, N_i) encoder) is connected to C_j (an (n_j, N_j) encoder) by an interleaver of size N_j. We have the "boundary conditions" $N_j = n_i$ and $w_j = h_i$.

4. The Interleaving Gain Exponent Conjecture.

In this section we will consider systems of the form depicted in Figure 1, in which the individual encoders are truncated convolutional encoders, and study the behavior of the average ML decoder error probability as the input block length N approaches infinity. If $A_{w,h}^N$ denotes the IOWE when the input block has length N, we introduce the following notation for the union bound (2.2) for systems of this type:

$$P_W^{\mathrm{UB}} \stackrel{\mathrm{def}}{=} \sum_{h=1}^{n} \left(\sum_{w=1}^{N} A_{w,h}^N \right) z^h. \tag{4.1}$$

Next we define, for each fixed $w \geq 1$ and $h \geq 1$,

$$\alpha(w,h) = \limsup_{N\to\infty} \log_N A_{w,h}^N. \tag{4.2}$$

It follows from this definition that if w and h are fixed,

$$A_{w,h}^N z^h = O(N^{\alpha(w,h)+\epsilon}) \qquad \text{as } N \to \infty,$$

for any $\epsilon > 0$. Thus if we define

$$\beta_M = \max_{h\geq 1} \max_{w \geq 1} \alpha(w,h). \tag{4.3}$$

it follows that for all w and h,

$$A_{w,h}^N z^h = O(N^{\beta_M+\epsilon}) \qquad \text{as } N \to \infty,$$

for any $\epsilon > 0$. The parameter β_M, which we shall call the *interleaving gain exponent* (IGE), was first introduced in [2] and [3] for parallel concatenation and later in [4] for serial concatenation. Extensive numerical simulations, and theoretical considerations that are not fully rigorous lead to the following conjecture about the behavior of the union bound for systems of the type shown in Figure 1.

The IGE Conjecture. *There exists a positive number γ_0, which depends on the q component convolutional codes and the tree structure of the overall system, but not on N, such that for any fixed $E_n/N_0 > \gamma_0$,as the block length N becomes large,*

$$P_W^{\mathrm{UB}} = O(N^{\beta_M}) \tag{4.4}$$

Eq. (4.4) implies that if $\beta_M < 0$, then for a given $E_b/N_0 > \gamma_0$ the *word* error probability of the concatenated code decreases to zero as the input block size is increased. This is summarized by saying that there is *word error probability interleaving gain.*[1]

In [7], we discuss the calculation of $\alpha(w,h)$ and β_M for a concatenated system of the type depicted in Figure 1, using analytical tools introduced in [3] and [4]. For example, for the parallel concatenation of q codes, with $q-1$ interleavers, we have

$$\beta_M \leq -q+2,$$

with equality if and only if each of the component codes is recursive. For a "classical" turbo code with $q=2$, we have $\beta_M = 0$, so there is no word error probability interleaving gain. This suggests that the word error probability for classic turbo codes will not improve with input block size, which is in agreement with simulations.

[1] There is a similar conjecture for the bit error probability which we do not discuss in this paper. Suffice it to say that the interleaving gain exponent for bit error probability is $\beta_M - 1$.

As another example, consider the serial concatenation of two convolutional codes. If the inner code is recursive then,

$$\beta_M \leq -\left\lfloor \frac{d^{\mathrm{o}}_{\mathrm{free}}+1}{2} \right\rfloor + 1,$$

where $d^{\mathrm{o}}_{\mathrm{free}}$ is the minimum distance of the outer code. Therefore, for serial concatenated codes, if $d^o_f \geq 3$ there is interleaving gain for word error probability. (If the inner code is nonrecursive $\beta_M \geq 0$ and there is no interleaving gain.)

5. A Class of Simple Turbo-Like Codes.

In this section we will introduce a class of turbo-like codes which are simple enough so that we can prove the IGE conjecture. We call these codes *repeat and accumulate* (RA) codes. The general idea is shown in Figure 3. An information block of length N is repeated q times, scrambled by an interleaver of size qN, and then encoded by a rate 1 *accumulator*. The accumulator can be viewed as a truncated rate-1 recursive convolutional encoder with transfer function $1/(1+D)$, but we prefer to think of it as a block code whose input block $[x_1, \ldots, x_n]$ and output block $[y_1, \ldots, y_n]$ are related by the formula

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_1 + x_2 \\ y_3 &= x_1 + x_2 + x_3 \\ &\ \vdots \\ y_n &= x_1 + x_2 + x_3 + \cdots + x_n. \end{aligned} \tag{5.1}$$

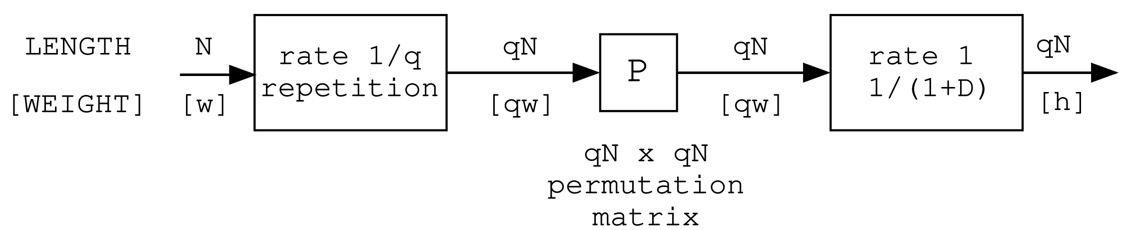

Figure 3. Encoder for a (qN, N) repeat and accumulate code. The numbers above the input-output lines indicate the length of the corresponding block, and those below the lines indicate the weight of the block.

To apply the union bound from Section 2 to the class of RA codes, we need the input-output weight enumerators for both the (qn, n) repetition code, and the (n, n) accumulator code. The outer repetition code is trivial: if the input block has length n, we have

$$A^{(o)}_{w,h} = \begin{cases} 0 & \text{if } h \neq qw \\ \binom{n}{w} & \text{if } h = qw. \end{cases} \tag{5.2}$$

The inner accumulator code is less trivial, but it is possible to show that (again assuming the input block has length n):

$$A^{(i)}_{w,h} = \binom{n-h}{\lfloor w/2 \rfloor}\binom{h-1}{\lceil w/2 \rceil - 1}. \tag{5.3}$$

It follows then from the general formula (3.1), that for the (qN, N) RA code represented by Figure 3, the ensemble IOWE is

$$\begin{aligned} A^{(N)}_{w,h} &= \sum_{h_1=0}^{qN} \frac{A^{(o)}_{w,h_1} A^{(i)}_{h_1,h}}{\binom{qN}{qw}} \\ &= \frac{\binom{N}{w}\binom{qN-h}{\lfloor qw/2 \rfloor}\binom{h-1}{\lceil qw/2 \rceil - 1}}{\binom{qN}{qw}}. \end{aligned} \tag{5.4}$$

From (5.4) it is easy to compute the parameters $\alpha(w,h)$ and β_M in (4.2) and (4.3). The result is

$$\alpha(w,h) = -\left\lceil \frac{(q-2)w}{2} \right\rceil \tag{5.5}$$

$$\beta_M = -\left\lceil \frac{(q-2)}{2} \right\rceil. \tag{5.6}$$

It follows from (5.6) that an RA code can have word error probability interleaving gain only if $q \geq 3$.

We are now prepared to use the union bound to prove the IGE conjecture for RA codes. In order to simplify the exposition as much as possible, we will assume for the rest of this section that $q = 4$, the extension to arbitrary $q \geq 3$ being straightforward but rather lengthy. For $q = 4$, (5.6) becomes $\beta_M = -1$, so the IGE conjecture is $P_W^{\mathrm{UB}} = O(N^{-1})$ for $E_b/N_0 > \gamma_0$ in this instance.

The union bound (2.2) for the ensemble of $q = 4$ RA codes is, because of (5.4),

$$P_W^{\mathrm{UB}} = \sum_{h=2}^{4N} \sum_{w=1}^{h/2} \frac{\binom{N}{w}\binom{4N-h}{2w}\binom{h-1}{2w-1}}{\binom{4N}{4w}} z^h. \tag{5.7}$$

Denote the (w,h)th term in the sum (5.7) by $T_N(w,h)$:

$$T_N(w,h) \stackrel{\mathrm{def}}{=} A_{w,h} z^h = \frac{\binom{N}{w}\binom{4N-h}{2w}\binom{h-1}{2w-1}}{\binom{4N}{4w}} z^h.$$

Using standard techniques (e.g. [8, Appendix A]), it is possible to show that for all (w,h),

$$T_N(w,h) \leq D 2^{h[F(x,y) + \log_2 z]}, \tag{5.8}$$

where $D = 4/\sqrt{\pi}$ is a constant, $x = w/4N$, $y = h/4N$,

$$F(x,y) = \frac{-\frac{3}{4}H_2(4x) + (1-y)H_2(\frac{2x}{1-y}) + yH_2(\frac{2x}{y})}{y},$$

and $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. The maximum of the function $F(x,y)$ in the range $0 \le 2x \le y \le 1-2x$ occurs at $(x,y) = (0.100, 0.371)$ and is 0.562281, so that if $\log_2 z < -0.562281$, the exponent in (5.8) will be negative.

Let us therefore assume that $\log_2 z < -0.562281$, which is equivalent to $E_b/N_0 = -(1/r)\ln z = -4\ln z \ge 4 \cdot \ln 2 \cdot 0.562281 = 1.559 = 1.928$ dB. If E is defined to be $E = -\log_2 z + 0.562281$, it follows from (5.8) for all w and h,

$$T_N(w,h) \le D2^{-hE}. \tag{5.9}$$

What (5.9) tells us is that if $E_b/N_0 > 1.928$ dB, most of the terms in the union bound (5.7) will tend to zero rapidly, as $N \to \infty$. The next step in the proof is to break the sum in (5.7) into two parts, corresponding to those terms for which (5.9) is helpful, and those for which it is not. To this end, define

$$h_N \stackrel{\text{def}}{=} \frac{3}{E}\log_2 N,$$

and write

$$\begin{aligned} P_W^{\text{UB}} &= \sum_{h=2}^{4N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= \sum_{h=2}^{h_N}\sum_{w=1}^{h/2} T_N(w,h) + \sum_{h=h_N+1}^{4N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= S_1 + S_2. \end{aligned}$$

It's easy to verify that when N is large enough, $A_{w+1,h}/A_{w,h} < 1$ for $h \le h_N$ and $w \le h/2 \le h_N/2$, which shows $A_{w,h}$ is a decreasing function of w for large N. Thus the sum S_1 can be overbounded as follows (we omit some details):

$$\begin{aligned} S_1 &= \sum_{h=2}^{h_N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= \sum_{h=2}^{h_N} T_N(1,h) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} T_N(w,h) \\ &= O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} T_N(w,h) \\ &\le O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} A_{2,h} z^h \\ &= O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} O(h^3/N^2) z^h \\ &= O(N^{-1}) + O(h_N^5/N^2) \\ &= O(N^{-1}). \end{aligned}$$

For the sum S_2, we bound each term $T_N(w,h)$ by (5.9):

$$\begin{aligned} S_2 &= \sum_{h=h_N+1}^{4N} \sum_{w=1}^{h/2} T_N(w,h) \\ &\le \sum_{h_N+1}^{4N} \sum_{w=1}^{h/2} D 2^{-hE} \\ &= D/2 \sum_{h_N+1}^{4N} h 2^{-hE} \\ &\le D \frac{2^{-E h_N}(h_N+1)}{(1-2^{-E})^2} \\ &= O(N^{-3}\log_2 N) \\ &= o(N^{-2}). \end{aligned}$$

We have therefore shown that for the ensemble of $q = 4$ RA codes, if $E_b/N_0 > 1.928$ dB,

$$P_W^{\mathrm{UB}} = S_1 + S_2 = O(N^{-1}) + o(N^{-1}) = O(N^{-1}), \tag{5.10}$$

which as we saw above, is the IGE conjecture in this case.

Although the union bound gives a proof of the IGE conjecture for RA codes, the resulting value of γ_0 is by no means the best possible. Indeed, if we use the recent Viterbi-Viterbi improved union bound [9] to bound the sum S_2, we can lower the value of γ_0 considerably, e.g. for $q = 4$ from 1.928 dB to 0.313 dB. In Figure 4 and Table 1 we display our numerical results on RA codes. There we compare the "cutoff threshold" γ_0 for RA codes with q in the range $3 \le q \le 8$ using both the classical union bound and the Viterbi-Viterbi improved union bound to the cutoff threshold for the ensemble of all codes (i.e., "random codes") of a fixed rate. We believe that these values of γ_0 can be reduced still further, for example by using the bound of [6] instead of the Viterbi-Viterbi bound.

q	3	4	5	6	7	8
RA Codes (Union Bound)	2.200	1.928	1.798	1.721	1.670	1.631
Random Codes (Union Bound)	2.031	1.853	1.775	1.694	1.651	1.620
RA Codes (Viterbi Bound)	1.112	0.313	−0.125	−0.402	−0.592	−0.731
Random Codes (Viterbi Bound)	0.214	−0.224	−0.486	−0.662	−0.789	−0.885
Binary Shannon Limit	−0.495	−0.794	−0.963	−1.071	−1.150	−1.210

Table 1. Numerical data gleaned from Figure 4.

6. Performance of RA Codes with Iterative Decoding.

The results of this paper show that the performance of RA codes *with maximum-likelihood decoding* is very good. However, the complexity of ML decoding of RA

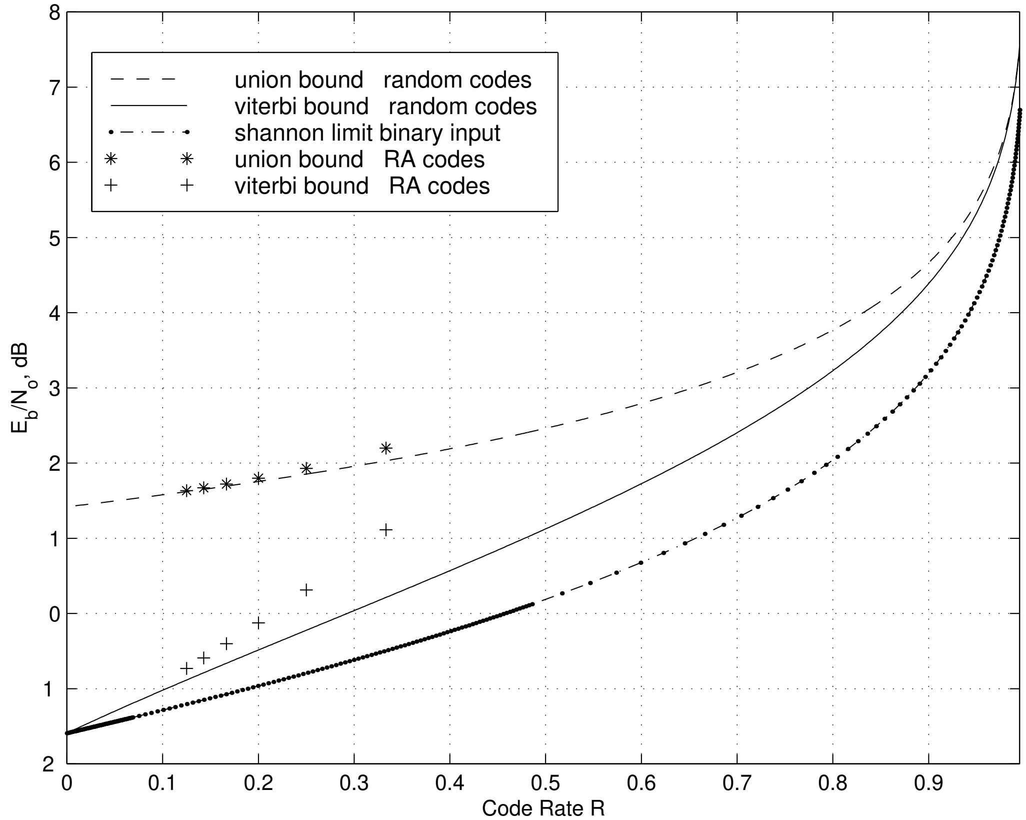

Figure 4. Comparing the RA code "cutoff threshold" to the cutoff rate of random codes using both the classical union bound and the Viterbi-Viterbi improved union bound.

codes, like that of all turbo-like codes, is prohibitively large. But an important feature of turbo-like codes is the availability of a simple iterative, message passing decoding algorithm that approximates ML decoding. We wrote a computer program to implement this "turbo-like" decoding for RA codes with $q = 3$ (rate 1/3) and $q = 4$ (rate 1/4), and the results are shown in Figure 5. We see in Figure 4, for example, that the empirical cutoff threshold for RA codes for $q = 3$ appears to be less than 1 dB, compared to the upper bound of 1.112 dB found in Table 1.

References.

1. C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo codes," *Proc. 1993 IEEE International Conference on Communications*, Geneva, Switzerland (May 1993), pp. 1064–1070.
2. S. Benedetto and G. Montorsi, "Unveiling turbo codes: some results on parallel concatenated coding schemes", *IEEE Trans. on Inf. Theory*, vol. 42, no. 2 (March 1996), pp. 409–428..
3. S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Transactions on Communications*, vol. 44, no. 5, (May 1996) pp. 591–600.
4. S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. on Information Theory*, vol. 44, no. 3, (May 1998), pp. 909–926.

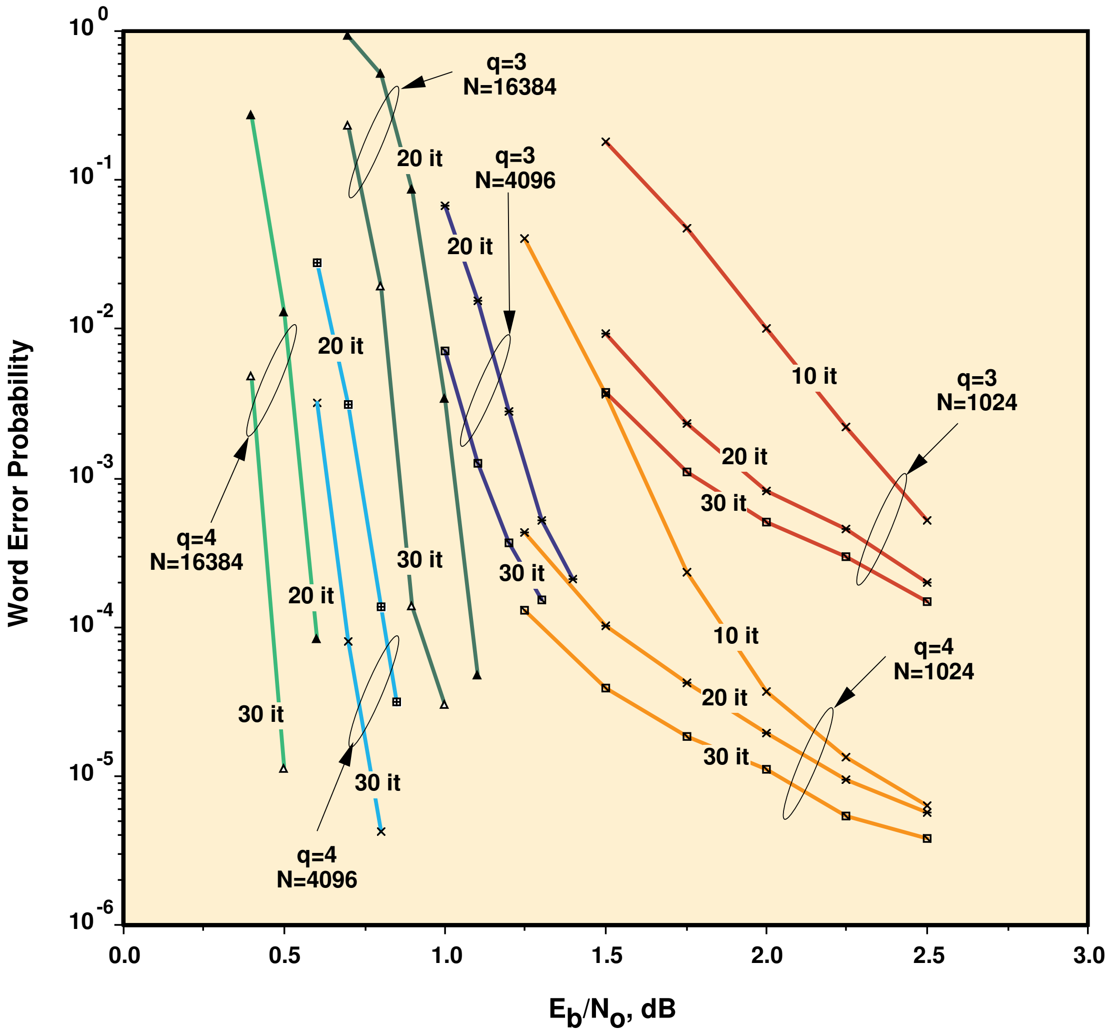

Figure 5. Simulated performance of iterative decoding of RA codes on an AWGN channel.

5. D. Divsalar and F. Pollara, "On the design of turbo codes," *TDA Progress Report* vol. 42-123 (November 15, 1995), pp. 99–121.
6. S. Dolinar, L. Ekroot, and F. Pollara, "Improved Error Probability Bounds for Block Codes for the Gaussian channel," Proc. 1994 ISIT, p. 243.
7. D. Divsalar and R. J. McEliece, "On the design of concatenated coding systems with interleavers," JPL TMO Progress Report vol. 2-134 (August 15, 1998), pp. 1–22. (`http://tmo.jpl.nasa.gov/tmo/progress_report/42-134/134D.pdf` .)
8. W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes, 2nd. ed.* Cambridge, Mass.: The MIT Press, 1972.
9. A. J. Viterbi and A. M. Viterbi, "Improved union bound on linear codes for the input-binary AWGN channel, with applications to turbo decoding," Proc. Winter 1998 Information Theory Workshop, San Diego California, Feb. 1998, p. 72.