Upper bounds to error probabilities of coded systems beyond the cutoff rate

Dariush Divsalar* Ezio Biglieri†

June 30, 2003

Abstract

A family of upper bounds to error probabilities of coded systems was recently proposed by Divsalar [5]. These bounds are valid for transmission over the additive white Gaussian noise channel, and require only the knowledge of the weight spectrum of the code words. After illustrating these bounds, we extend them to fading channels. Contrary to the union bound, our bounds maintain their effectiveness below the signal-to-noise ratio at which the cutoff rate of the channel equals the rate of the code. Some applications are shown: first, we derive upper bounds to the minimum signal-to-noise ratio necessary to achieve zero error probability as the code block length increases to infinity. Next, we use our bounds to predict the performance of turbo codes and low-density parity-check codes.

1 Introduction and motivation of the work

During the years, much effort has been spent in the search for close approximations to the error probability of systems in which coding is used in conjunction with maximum-likelihood decoding (here we are especially interested in linear binary codes, so we shall restrict our attention to these without any further stipulation). In many instances the union bound provides a useful tool for the prediction of system performance at intermediate-to-high signal-to-noise ratios (SNR). It is easy to compute, and requires only the knowledge of the weight spectrum of the code; however, it becomes looser and looser as the SNR decreases, to the point of being totally useless when the SNR approaches the value at which the cutoff rate R_0 of the channel equals the code rate R_c.

*Jet Propulsion Laboratory – Caltech USA. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

†Politecnico di Torino, Italy

The recent discovery of an easily decodable family of codes with good error properties even beyond the channel cutoff rate, and close to capacity [4], has rekindled the interest in bounds that overcome the R_0-limitation of the union bound, while keeping the upsides of it. Specifically, these new bounds should be easily computed, and depend only on the weight spectrum of the code: the latter property is especially important in view of the fact that with turbo codes only the weights, averaged with respect to the possible choices of the interleaver, are usually available.

For recent work in this area, see, for example, [7, 8, 14, 15, 16, 17, 18, 19, 20, 22, 23]. Of late, one such family of upper bounds was proposed in [5]. This family turns out to yield the tightest known approximation to the error probability of turbo codes with large block lengths. Here we first elaborate on these bounds, by showing how they can be further tightened through riddance of a number of unnecessary terms in a summation. Further, we extend these bounds to the fading channel. This paper is organized as follows. In Section 2 we briefly discuss the usefulness of error-probability bounds, and define the "critical rate" of a bound. In Section 3 we expound the new bounds for the additive Gaussian noise (AWGN) channel and their improvements. In Section 4 we derive bounds for the fading channel and show their application to the performance prediction of turbo codes and of low-density parity-check codes.

2 Range of usefulness of error-probability bounds

Consider a generic bound on word-error probability of a binary code with block length n, rate R_c bit/symbol (and hence $\ln 2 \cdot R_c$ nat/symbol), and minimum Hamming distance $d_{\min}$, transmitted over the AWGN channel with a ratio of energy-per-bit to noise-power spectral density $\eta \triangleq E_b/N_0$. Let the bound be expressed in the form

$$P(e) < \sum_{d=d_{\min}}^{n} e^{-nE(R_c,\eta,d)}. \tag{1}$$

where d denote the Hamming weights of the code. The exponent $E(R_c,\ \eta, d)$ implicitly depends on A_d the number of code words with weight d (the set $\{A_d\}$ is the *weight spectrum* of the code). We say that this bound is useful for a given η if $E(R_c,\ \eta, d) > 0$ for all d. For a given code rate, and d, let us consider $\min_\eta E(R_c,\ \eta, d) \geq 0$. Denote the minimizing η by $\eta_0(d)$. Then we define the critical signal-to-noise ratio η_{crit} as the value of η such that $\eta_{\text{crit}} = \max_d \eta_0(d)$ and the range of usefulness of the bound is $\eta > \eta_{\text{crit}}$.

Random codes. Consider for example random codes. It is well-known (see, e.g., [11,],[21, p. 138 ff.]) that a random block code of rate R_c and length n, transmitted over a binary-input, output-symmetric channel, has an error probability with maximum-likelihood decoding which is bounded (the "union-Bhattacharyya" bound [21, p. 129]), as

$$P(e) < e^{-n(R_0 - \ln 2 \cdot R_c)} \tag{2}$$

where R_0, the "cutoff rate" of the channel, has, for the AWGN channel, the expression

$$R_0 = \ln 2 - \ln\left(1 + e^{-R_c \eta}\right) \qquad \text{nat/symbol} \tag{3}$$

The bound (2) is useful for $\ln 2 \cdot R_c < R_0$, which yields the following critical SNR:

$$\eta_{\text{crit}} = -\frac{1}{R_c} \ln\left(2^{1-R_c} - 1\right) \tag{4}$$

Usefulness of the union bound for a specific code. Consider a linear binary block code, and the union-Bhattacharyya bound [21, p. 142]

$$P(e) < \sum_{d=d_{\min}}^{n} A_d e^{-dR_c\eta} \tag{5}$$

The bound based on (5) has the form (1), with

$$E(R_c, \eta, d) = -\frac{\ln A_d}{n} + \frac{d}{n} R_c \eta$$

and hence

$$\eta_{\text{crit}} = \max_d \eta_0(d) = \max_d \frac{1}{R_c} \frac{\ln A_d/n}{d/n} \tag{6}$$

The relation between (6) and (4) can be elucidated by picking a linear binary random code with block length n and 2^k code words (and hence rate $R_c = k/n$). This yields

$$A_d = \frac{1}{2^{n-k}}\binom{n}{d}$$

which, as $n \to \infty$, tends to $2^{-(n-k)}e^{nH(d/n)}$, with $H(\cdot)$ the binary entropy function

$$H(x) \triangleq -x \ln x - (1-x)\ln(1-x). \tag{7}$$

In conclusion, as $n \to \infty$ the maximum value of η_{crit} in (6) is achieved for $d/n = 1 - 2^{-(1-R_c)}$, and equals (4).

3 AWGN channel bounds

Consider in general the transmission of a geometrically-uniform [9] signal constellation $\mathcal{X}$, with $|\mathcal{X}| = M$, over the AWGN channel, modelled in the form

$$\mathbf{y} = \gamma\mathbf{x} + \mathbf{n} \tag{8}$$

where $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{n}$ are n-dimensional real vectors; in particular, $\mathbf{x} \in \mathcal{X}$ denotes the transmitted signal vector, $\mathbf{y}$ the received vector, $\mathbf{n}$ a random noise vector whose components are Gaussian random variables with mean zero and common variance 1, and γ is a known constant. We also assume that the code word components take on values ± 1, so that all signal vectors have equal energy $\| \mathbf{x} \|^2 = n$. With maximum-likelihood (ML) decoding, the word error probability when $\mathbf{x}$ is transmitted does not depend on $\mathbf{x}$ due to our assumption of a geometrically-uniform constellation. It can be written in the form

$$P(e) = \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \neq \mathbf{x}} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \tag{9}$$

where $\{\mathbf{x} \to \widehat{\mathbf{x}}\}$ denote the "pairwise error event." This is the probability that when $\mathbf{x}$ is transmitted the Euclidean distance (ML decoding metric) between the received vector $\mathbf{y}$ and $\widehat{\mathbf{x}}$ is smaller that the distance between $\mathbf{y}$ and $\mathbf{x}$, that is, $\{\mathbf{x} \to \widehat{\mathbf{x}}\} \triangleq \{\mathbf{y} \ : \| \mathbf{y} - \widehat{\mathbf{x}} \| < \| \mathbf{y} - \mathbf{x} \|\}$. Notice that the set of $\mathbf{y}$ such that $\{\mathbf{x} \to \widehat{\mathbf{x}}\}$ occurs is a half-space in $\mathbb{R}^n$, the locus of the points whose distance from $\mathbf{x}$ exceeds the distance from $\widehat{\mathbf{x}}$.

The number of terms in the union of (9) can be reduced if we remove any pairwise error event that can be written as a union of other events that are retained. In particular, the number of terms is minimized if we keep only the signal vectors $\mathbf{x}$ that are defined as follows. Define the Voronoi region of $\mathbf{x}$ as the set of vectors in the Euclidean n-dimensional space $\mathbb{R}^n$ that are closest to $\mathbf{x}$ than to any other $\widehat{\mathbf{x}}$. The Voronoi region is a convex polytope in $\mathbb{R}^n$, the intersection of the half-spaces described above. The number of facets in this polytope is usually much lower that $|\mathcal{X}|$: for this reason it is convenient to remove the redundancy and redefine $\mathcal{V}$ by using only the inequalities that are strictly necessary. To do this, we define the set of (Voronoi) *neighbors* of $\mathbf{x}$ as the minimal set $\mathcal{N}$ such that

$$\mathcal{V} = \{\mathbf{y} \in \mathbb{R}^n \ : \| \mathbf{y} - \mathbf{x} \| \leq \| \mathbf{y} - \widehat{\mathbf{x}} \|, \ \forall \widehat{\mathbf{x}} \in \mathcal{N}\} \tag{10}$$

The vector $\mathbf{x}$ itself does not belong to $\mathcal{N}$. Based on this definition of $\mathcal{N}$, we can rewrite the error probability in the form

$$P(e) = \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \tag{11}$$

Now, the union includes only $|\mathcal{N}|$ terms, the minimum possible number. . For future reference, we can derive from (11) a "minimal" union bound by writing

$$P(e) \leq \sum_{\widehat{\mathbf{x}} \in \mathcal{N}} \mathbb{P}\{\mathbf{x} \to \widehat{\mathbf{x}}\} \tag{12}$$

A bound tighter than (12) can be obtained as follows. Let d denote the generic Hamming distance of $\widehat{\mathbf{x}}$ from $\mathbf{x}$. Partition $\mathcal{N}$ into equivalence classes of vectors with the same value of d. Denote these by $\mathcal{N}_d$, and write

$$\begin{aligned} P(e) &= \mathbb{P}\left[\bigcup_{d \in \mathcal{D}} \bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_d} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \\ &\leq \sum_{d \in \mathcal{D}} \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_d} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \end{aligned}$$

$$= \sum_{d \in \mathcal{D}} \mathbb{P}[e_d] \tag{13}$$

where $\mathcal{D}$ is the set of distances from $\mathbf{x}$ of the vectors in $\mathcal{X}$, and

$$e_d \triangleq \bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_d} \{\mathbf{x} \to \widehat{\mathbf{x}}\} \tag{14}$$

is the event that, when $\mathbf{x}$ is transmitted, at least one $\widehat{\mathbf{x}}$ at distance d is nearer to $\mathbf{y}$ than $\mathbf{x}$.

3.1 Characterizing $\mathcal{N}$

The characterization of $\mathcal{N}$ is a well-studied problem (see, e.g., [1, 2, 3, 13]). Although it generally requires a knowledge of the code structure that extends beyond its weight spectrum, useful bounds are available. Let $\mathcal{X}$ be a binary linear block code with parameters n, k, $d_{\min}$. If w denotes the weight of the code word $\widehat{\mathbf{x}}$ (i.e., the number of -1s in it), then the following theorem [2] yields simple bounds to $\mathcal{N}$.

Theorem. For any binary linear block code

$$\{\widehat{\mathbf{x}} \in \mathcal{X} \ : \ 1 \le w \le 2d_{\min} - 1\} \subseteq \mathcal{N} \subseteq \{\widehat{\mathbf{x}} \in \mathcal{X} \ : \ 1 \le w \le n - k + 1\} \tag{15}$$

Moreover, if the weight $w \neq 0$ of $\widehat{\mathbf{x}}$ cannot be written as $w = i + j$, where $i \ge 1$, $j \ge 1$, and i, j are actual weights of words of $\mathcal{X}$, then $\widehat{\mathbf{x}} \in \mathcal{N}$. □

As an immediate consequence of (15), for a linear block code we can rewrite (13) as

$$P(e) \le \sum_{d=d_{\min}}^{n-k+1} \mathbb{P}[e_d] \tag{16}$$

The bound above can be further tightened by using the rest of the Theorem. An algorithm is also available [1] to derive the elements of $\mathcal{N}$ if the code words can be listed. For example, from tables in [1] we can infer that $\mathcal{N}$ for the $(31, 21)$ BCH code contains 107,198 words, while $|\mathcal{X}| = 2{,}097{,}152$. The number of words expurgated from the computation of the bound may consequently be very large, especially when the rate of the code is greater than $1/2$ (see [2]).

We should also observe that the words excluded have large distances (in fact, from the Theorem we see that all words not in $\mathcal{N}$ have a Hamming weight $w \geq 2d_{\min}$): thus, the effect of this expurgation would be especially felt at low signal-to-noise ratios.

3.2 The new bound

We now compute an upper bound to $P(e)$ based on (13). To do this, we use a technique advocated by Gallager in [10] and express $\mathbb{P}[e_d]$ as

$$\mathbb{P}[e_d] = \mathbb{P}[e_d, \mathbf{y} \in \mathcal{R}] + \mathbb{P}[e_d, \mathbf{y} \notin \mathcal{R}] \tag{17}$$

where $\mathcal{R}$ is any region in $\mathbb{R}^n$. Further, observe that

$$\mathbb{P}[e_d, \mathbf{y} \notin \mathcal{R}] \leq \mathbb{P}[\mathbf{y} \notin \mathcal{R}] \tag{18}$$

where for the RHS to be a good approximation of the LHS one should choose $\mathcal{R}$ in such a way that the two regions $\overline{\mathcal{R}}$ (the complement of $\mathcal{R}$ in $\mathbb{R}^n$) and $\{\mathbf{y} : \bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_d} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\}$ have about the same shape and size. In practice, the selection of $\mathcal{R}$ should be guided by computational simplicity. The new bound is based on the choice for $\mathcal{R}$ of an n-dimensional hypersphere centered at $\gamma\epsilon\mathbf{x}$ and with radius $\sqrt{n}R$. The parameters ϵ and R will be selected so as to obtain the tightest possible bound (notice that in general the choice of ϵ and R will depend on d).

By combining (18) and (17) we obtain the upper bound

$$\mathbb{P}[e_d] \leq \mathbb{P}[e_d, \mathbf{y} \in \mathcal{R}] + \mathbb{P}[\mathbf{y} \notin \mathcal{R}] \tag{19}$$

3.2.1 Computation of $\mathbb{P}[\mathbf{y} \notin \mathcal{R}]$

Let us compute the second term in the RHS of (19) first. We have

$$\mathbb{P}[\mathbf{y} \notin \mathcal{R}] = \mathbb{P}\left[\sum_{k=1}^{n} (y_k - \gamma\epsilon x_k)^2 \geq nR^2\right] = \mathbb{P}[W \geq 0] \tag{20}$$

where we have defined the RV $W \triangleq \sum_{k=1}^{n} (y_k - \gamma\epsilon x_k)^2 - nR^2$. We bound (20) using the Chernoff bound $\mathbb{P}[W \geq 0] \leq \mathbb{E}[e^{sW}]$, with $0 \leq s \leq \frac{1}{2}$. We observe that, under the assumption that $\mathbf{x}$ was transmitted, we have $\mathbf{y} = \gamma\mathbf{x} + \mathbf{n}$. Moreover, the components of $\mathbf{n}$ are independent RVs with mean zero and variance 1, and $\| \mathbf{x} \|^2 = n$. Thus, we obtain

$$\begin{aligned}\mathbb{E}[e^{sW}] &= e^{-snR^2} \prod_{k=1}^{n} \mathbb{E}\left[e^{s((1-\epsilon)\gamma x_k + n_k)^2}\right] \\ &= e^{-snR^2} \prod_{k=1}^{n} \frac{1}{\sqrt{1-2s}} e^{(1-\epsilon)^2\gamma^2 x_k^2 s/(1-2s)} \\ &= e^{-snR^2} g^n(\gamma, s, \epsilon) \qquad (21)\end{aligned}$$

where

$$g(\gamma, s, \epsilon) \triangleq \frac{1}{\sqrt{1-2s}} e^{(1-\epsilon)^2\gamma^2 s/(1-2s)} \qquad (22)$$

3.2.2 Computation of $\mathbb{P}[e_d,\ \mathbf{y} \in \mathcal{R}]$

We have

$$\begin{aligned}\mathbb{P}[e_d,\ \mathbf{y} \in \mathcal{R}] &= \mathbb{P}\left[\bigcup_{\hat{\mathbf{x}} \in \mathcal{N}_d} \{\mathbf{x} \to \hat{\mathbf{x}}\},\ \mathbf{y} \in \mathcal{R}\right] \qquad (23) \\ &\leq \sum_{\hat{\mathbf{x}} \in \mathcal{N}_d} \mathbb{P}\left[\{\mathbf{x} \to \hat{\mathbf{x}}\},\ \mathbf{y} \in \mathcal{R}\right] \\ &= \sum_{\hat{\mathbf{x}} \in \mathcal{N}_d} \mathbb{P}\left[\| \mathbf{y} - \hat{\mathbf{x}} \| \leq \| \mathbf{y} - \mathbf{x} \|,\ \mathbf{y} \in \mathcal{R}\right] \\ &= \sum_{\hat{\mathbf{x}} \in \mathcal{N}_d} \mathbb{P}\left[(\mathbf{y}, \mathbf{x}) \leq (\mathbf{y}, \hat{\mathbf{x}}),\ \| \mathbf{y} - \gamma\epsilon\mathbf{x} \|^2 \leq nR^2\right] \qquad (24)\end{aligned}$$

The corresponding Chernoff bound can be computed [10]:

$$\mathbb{P}[Z \geq 0,\ W \leq 0] \leq \mathbb{E}\left[e^{tZ + rW}\right], \qquad t \geq 0,\ r \leq 0 \qquad (25)$$

and with optimized $t = (1 - 2r\epsilon)\gamma/2$ has the form

$$\mathbb{P}[e_d,\ \mathbf{y} \in \mathcal{R}] \leq A_d e^{-nrR^2} f^d(\gamma,\ r,\ \epsilon) g^{n-d}(\gamma,\ r,\ \epsilon) \qquad (26)$$

where $g(\cdot, \cdot, \cdot)$ was defined in (22), and

$$f(\gamma, r, \epsilon) \triangleq \frac{1}{\sqrt{1-2r}} e^{-\frac{\gamma^2}{2}(1-2r\epsilon^2)} \qquad (27)$$

If we select the value of e^{nR^2} that minimizes the Chernoff bound, and we define for notational simplicity

$$A(\gamma,\, r,\, \epsilon) \triangleq A_d f^d(\gamma,\, r,\, \epsilon) g^{n-d}(\gamma,\, r, \epsilon) \tag{28}$$

$$B(\gamma,\, s,\, \epsilon) \triangleq g^n(\gamma,\, s\,, \epsilon) \tag{29}$$

we obtain the neat bound

$$\mathbb{P}[e_d] \leq e^{H(\frac{s}{s-r})} A^{\frac{s}{s-r}}(\gamma,\, s,\, \epsilon) B^{-\frac{r}{s-r}}(\gamma,\, r,\, \epsilon) \tag{30}$$

where $H(.)$ was defined by (7). Let $2s = (\beta-\rho)/(1-\rho)$, and $2r = (\rho-\beta)/\rho$ with $0 \leq \rho \leq \beta \leq 1$. As discussed in the Appendix, the factor $e^{H(\rho)}$ (whose value ranges between 1 and 2) can be ignored in the bound. Some algebra is required to minimize the bound (30) with respect to ϵ, ρ, and β. First the optimum $\epsilon = (\beta - \rho(1 - \frac{d}{n}(1-\beta)))/((1-\frac{d}{n}(1-\beta))(\beta-\rho))$ is obtained. Next we obtain the optimum ρ as

$$\rho^* \triangleq \frac{1}{1 + \dfrac{1-\beta}{\beta} e^{2(\ln A_d)/n}} \tag{31}$$

Then the bound can be expressed as

$$\mathbb{P}[e_d] \leq \exp\{-nE(\gamma^2/2,\, d/n,\, \beta)\} \tag{32}$$

where

$$E(\gamma^2/2,\, d/n,\, \beta) \triangleq \frac{1}{2}\ln(1-\beta+\beta e^{-2(\ln A_d)/n}) + \frac{\beta d/n}{1-(1-\beta)d/n}\frac{\gamma^2}{2} \tag{33}$$

and the optimum β is

$$\beta \triangleq \frac{1-d/n}{d/n}\left[\sqrt{\frac{\gamma^2}{2}\frac{2d/n}{1-d/n}\frac{1}{1-e^{-2(\ln A_d)/n}} + \left(1+\frac{\gamma^2}{2}\right)^2 - 1} - \left(1+\frac{\gamma^2}{2}\right)\right] \tag{34}$$

Since $0 \leq \beta \leq 1$, the bound is valid for $(1-e^{-2\ln A_d/n})\frac{1-d/n}{2d/n} \leq \frac{\gamma^2}{2} \leq \frac{e^{2(\ln A_d)/n}-1}{2(1-d/n)d/n}$ for $\frac{\gamma^2}{2} \geq \frac{e^{2(\ln A_d)/n}-1}{2(1-d/n)d/n}$ we set $\beta = 1$. The important byproduct of this bound is the tightest closed form critical signal-to noise ratio η_{crit} that can be obtained as

$$\eta_{\text{crit}} = \max_d \frac{1}{R_c}(1-e^{-2\ln A_d/n})\frac{1-d/n}{2d/n} \tag{35}$$

A discussion of the relations between this new bound, the union bound, and other previously derived bounds can be found in [5]. In particular, the new bound has the smallest critical SNR among the closed-form bounds known to the authors. Specifically, its critical SNR is the same as for the bounds in [7, 14, 8], and is smaller than that of the bounds in [22, 12]. Indeed we can show η_{crit} (35) can be bounded as

$$\eta_{\text{crit}} \leq \max_d \frac{1}{R_c}(1 - e^{-2\ln A_d/n})\frac{1}{2d/n} \triangleq \eta_{\text{crit}}(\text{Hughes}) \tag{36}$$

or can be bounded as

$$\eta_{\text{crit}} \leq \max_d \frac{1}{R_c}(\ln A_d/n)\frac{1}{d/n} \triangleq \eta_{\text{crit}}(\text{Viterbi}) \tag{37}$$

where the upperbound on η_{crit} can be derived using error probability bounds by Hughes [12] and Viterbi-Viterbi [22] respectively. The above bounds on η_{crit} further can be upper-bounded by (6). In [5] it is shown how (32) can be extended to the computation of bit error probability by simply replacing A_d with $\sum_{i=1}^{R_c n} \frac{i}{R_c n} A_{i,d}$ where $A_{i,d}$ is the number of codewords with input weight i and output weight d.

3.3 Refining the bound

Observe that the bound based on (32) is especially tight at low signal-to-noise ratios; since at higher values of SNR the bound $\mathbb{P}[e_d] \leq A_d Q(\sqrt{\gamma^2 d})$ may be tighter. Using the above two facts, we can refine (32) in the form

$$\mathbb{P}[e_d] \leq \min[\exp\{-nE(\gamma^2/2, d/n, \beta)\}, A_d Q(\sqrt{\gamma^2 d})] \tag{38}$$

3.4 Comparison with other bounds

A comparison of the bound derived here with those obtained previously can be obtained as follows. We compute the minimum achievable signal-to-noise ratio E_b/N_0 such that the error probability bound for a given code tends to zero as the code block length increases to infinity.

Fig. 1 shows such a comparison for random codes and Repeat-Accumulate (RA) codes (details of the calculations can be found in [5]).

4 Independent fading channel bounds

The channel model here is described by the equation

$$\mathbf{y} = \gamma \boldsymbol{\alpha} \mathbf{x} + \mathbf{n} \tag{39}$$

where $\boldsymbol{\alpha} = \text{diag}\,(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the diagonal matrix of the fading gains affecting the components of the transmitted vector $\mathbf{x}$. The fading gains α_i are independent, identically distributed random variables which we shall assume each to be Rayleigh-distributed with $\mathbb{E}[\alpha_i^2] = 1$, i.e., to have the probability density function $f_\alpha(r) = 2re^{-r^2}$, with $r \geq 0$. We use here the upper bound, derived from (19),

$$\mathbb{P}[e_d] = \mathbb{E}_{\boldsymbol{\alpha}} \mathbb{P}[e_d \mid \boldsymbol{\alpha}] \leq \mathbb{E}_{\boldsymbol{\alpha}} \mathbb{P}[e_d,\, \mathbf{y} \in \mathcal{R} \mid \boldsymbol{\alpha}] + \mathbb{E}_{\boldsymbol{\alpha}} \mathbb{P}[\mathbf{y} \notin \mathcal{R} \mid \boldsymbol{\alpha}] \tag{40}$$

4.1 Selection of the region $\mathcal{R}$

Notice first that the decision metric in this case is based on the minimization of the norm

$$\| \mathbf{y} - \gamma \boldsymbol{\alpha} \mathbf{x} \| \tag{41}$$

The simplest region suggested by (41) is a sphere with radius $\sqrt{n}R$ centered at $\epsilon\gamma\boldsymbol{\alpha}\mathbf{x}$, with ϵ and R parameters to be optimized. However, the resulting bound is not tight. Another choice for $\mathcal{R}$ consists of a sphere centered at $\epsilon\gamma\boldsymbol{\alpha}\mathbf{x}$ but the radius of sphere depends on $\boldsymbol{\alpha}$. A simple choice is

$$\mathcal{R} = \{\mathbf{y} \mid \| \mathbf{y} - \epsilon\gamma\boldsymbol{\alpha}\mathbf{x} \|^2 \leq nR^2 + \lambda\gamma^2 \| \boldsymbol{\alpha}\mathbf{x} \|^2\} \tag{42}$$

Notice that in this case equivalently we have

$$\mathcal{R} = \{\mathbf{y} \mid \ \zeta \| \mathbf{y} \|^2 - \gamma(\mathbf{y}, \boldsymbol{\alpha}\mathbf{x}) + \phi\gamma^2 \| \boldsymbol{\alpha}\mathbf{x} \|^2 \leq nR^2\} \tag{43}$$

where ζ, ϕ, and R are now the parameters to be optimized. In the following we present the bounds resulting from this choice of $\mathcal{R}$.

4.2 The new bound

By replicating the computations described in section 3.2, except that now for independent Rayleigh fading we should have

$$f(\gamma, r, \zeta, \phi) = \frac{1}{1+\frac{\gamma^2}{2}(1-2r\phi)} \frac{1}{\sqrt{1-2r\zeta}} \tag{44}$$

$$g(\gamma, x, \zeta, \phi) = \frac{1}{1+\frac{\gamma^2}{2}(1-2x\phi - \frac{(1-x)^2}{1-2x\zeta})} \frac{1}{\sqrt{1-2x\zeta}} \tag{45}$$

where x can be the parameter r or s (see (28), (29), and (30)). Now let $\rho \triangleq s/(s-r)$, then $s = -r\rho/(1-\rho)$ further define a new parameter $\beta \triangleq \rho(1-2r\zeta)$, where $0 \leq \rho \leq \beta \leq 1$. As for the AWGN bound, the factor $e^{H(\rho)}$ in (30) can be removed and the bound tightened. After these change of variables we obtain a new bound for independent Rayleigh fading in the form

$$\mathbb{P}[e_d] \leq \exp\{-nE(\gamma^2/2,\, d/n,\, \rho,\, \beta,\, r, \phi)\} \tag{46}$$

where, after some algebra,

$$\begin{aligned} E(\gamma^2/2,\, d/n,\, \rho,\, \beta,\, r, \phi) \;\triangleq\; & -\rho(\ln A_d)/n + \frac{\rho}{2}\ln\frac{\beta}{\rho} + \frac{1-\rho}{2}\ln\frac{1-\beta}{1-\rho} \\ & +\rho\frac{d}{n}\ln\left[1+\frac{\gamma^2}{2}(1-2r\phi)\right] \\ & +\rho(1-d/n)\ln\left[1+\frac{\gamma^2}{2}\left(1-2r\phi-\frac{(1-r)^2\rho}{\beta}\right)\right] \\ & +(1-\rho)\ln\left[1+\frac{\gamma^2}{2}\left(\frac{1-\rho(1-2r\phi)}{1-\rho} - \frac{(1-\rho(1-r))^2}{(1-\rho)(1-\beta)}\right)\right] \end{aligned} \tag{47}$$

This bound first appeared in [6] without proof. The bound should be minimized with respect to ϕ, ρ, β, and r. The minimum with respect to ϕ can be obtained in a closed form, while the remaining minimizations must be performed numerically. The bound can be further tightened at high SNR as we did in Section 3.3, by using

$$\mathbb{P}[e_d] \leq \min \left\{ \exp\{-nE(\gamma^2/2,\, d/n,\, \rho,\, \beta,\, r, \phi)\},\, A_d \frac{1}{\pi} \int_0^{\pi/2} \left[\frac{\sin^2\theta}{\sin^2\theta + \gamma^2/2} \right]^d d\theta \right\} \tag{48}$$

Usefulness of the new bound. The suboptimum choice $\rho = \beta = 1$, $r = 0$, and $\phi = 0.5$ yields

$$E(\gamma^2/2,\, d/n,\, 1,\, 1,\, 0,\, 0.5) = -\frac{\ln A_d}{n} + \frac{d}{n} \ln \left[1 + \frac{\gamma^2}{2} \right], \tag{49}$$

which is the exponent of the union-Bhattacharyya bound for the independent Rayleigh fading channel. By recalling the discussion in Section 2, as $n \to \infty$ this suboptimum bound becomes useless when

$$\ln \left[1 + \frac{\gamma^2}{2} \right] = \max_{d/n} \frac{(\ln A_d)/n}{d/n} \tag{50}$$

For large block length n and random codes we have

$$\frac{\ln A_d}{n} = H(d/n) - (1 - R_c) \ln 2 \tag{51}$$

so that

$$\max_{d/n} \frac{(\ln A_d)/n}{d/n} = -\ln \left[2^{1-R_c} - 1 \right] \tag{52}$$

and the minimum SNR for the validity of the bound turns out to be

$$\left(\frac{\gamma^2}{2} \right)_{\min} = \frac{1}{2^{1-R_c} - 1} - 1 \tag{53}$$

Since $\gamma^2/2 = R_c \eta$, this corresponds to

$$\eta_{\text{crit}} = \frac{1}{R_c(2^{1-R_c} - 1)} - \frac{1}{R_c} \tag{54}$$

which, not unexpectedly, is the equation for the cutoff rate of the independent Rayleigh fading channel. Since with optimum parameters our bound is tighter than the union-Bhattacharyya bound, we can expect it to have a lower η_{crit}. This is confirmed by numerical calculations, which show that for $R_c = 1/2$ the new bound is useful above 3.06 dB, while the cutoff rate is achieved at 4.52 dB. Thus, our bound outperforms the union bound by about 1.5 dB, although its usefulness cannot extend to capacity (which for a rate-1/2 code is 1.8 dB). As $R_c \to 0$, the union-bound critical SNR tends to 1.46 dB, while the new bound is useful up to -1 dB. This suggests that the new bound is tighter for low-rate codes.

4.3 Examples

The simple bound for AWGN channel can be used to obtain the word error probability of Low-Density Parity-Check (LDPC) codes with ML decoding. as shown in . The example of Fig. 2 refers to rate-$1/2$ (n, j, k) LDPC codes with $n = 10000$, $j = 3, 4, 5, 6$, and $k = 2j$. In the second example, shown in Fig. 3, the simple bound for AWGN channel is applied to obtain the ML performance of rate-$1/4$ RA codes. The figure also shows the simulated performance of the (suboptimum) iterative turbo decoder. Our third example, illustrated in Fig. 4, applies the simple bound for Rayleigh fading channel to obtain the ML performance of rate-$1/4$ RA codes. The simulated performance of the iterative turbo decoder is also shown.

5 Conclusions

We have presented a general bounding technique, obtained from Gallager's (19). By choosing a suitable region $\mathcal{R}$ in a parametric form and optimizing its parameters, a bound was obtained which is useful beyond the cutoff rate, and hence lends itself to be applied to coding schemes operating at low signal-to-noise ratios. After summarizing the application of our bounding technique to the AWGN channel, we have extended it to fading channels.

APPENDIX: A relation between two bounds due to Gallager

In this Appendix we prove that the factor $\exp H(\rho)$ can be omitted from bound (32). We do this by showing a relation between two upper bounds proposed by Gallager and used in several recent investigations into bounds that extend beyond the cutoff rate. These are (19), which will be referred to in the sequel as the *geometric bound* (GB), and the *generalized union bound* (GUB) [11], valid for any $\rho > 0$,

$$\mathbb{P}[e_d] \leq \sum_{\mathbf{y}} [P(\mathbf{y} \mid \mathbf{x})]^{1/(1+\rho)} \left\{ \sum_{\widehat{\mathbf{x}} \in \mathcal{X}_d \,:\, \widehat{\mathbf{x}} \neq \mathbf{x}} [P(\mathbf{y} \mid \widehat{\mathbf{x}})]^{1/(1+\rho)} \right\}^{\rho} \tag{55}$$

where $\mathcal{X}_d$ denotes the subset of code words of $\mathcal{X}$ whose Hamming distance from x is d.

First, rewrite (55) in the form

$$\mathbb{P}[e_d] \leq \sum_{\mathbf{y}} P(\mathbf{y} \mid \mathbf{x}) \left\{ \sum_{\widehat{\mathbf{x}} \in \mathcal{X}_d \,:\, \widehat{\mathbf{x}} \neq \mathbf{x}} \left[\frac{P(\mathbf{y} \mid \widehat{\mathbf{x}})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{\lambda} \right\}^{\rho} \tag{56}$$

where we write λ for $1/(1+\rho)$. Next choose, following Duman and Salehi [8], any nonnegative function $f(\mathbf{y})$, and rewrite the bound above in the equivalent form

$$\mathbb{P}[e_d] \leq \sum_{\mathbf{y}} f(\mathbf{y}) \left\{ \sum_{\widehat{\mathbf{x}} \in \mathcal{X}_d \,:\, \widehat{\mathbf{x}} \neq \mathbf{x}} \left[\frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{-1/\rho} \left[\frac{P(\mathbf{y} \mid \widehat{\mathbf{x}})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{\lambda} \right\}^{\rho} \tag{57}$$

Inequality (57) can be generalized (and possibly tightened) by introducing a new parameter $s \geq 0$ as follows:

$$\begin{aligned}
\mathbb{P}[e_d] \;\leq\; & \left\{ \sum_{\mathbf{y}} P(\mathbf{y} \mid \mathbf{x}) \left[\frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{s} \right\}^{1-\rho} \\
& \times \left\{ \sum_{\widehat{\mathbf{x}} \in \mathcal{X}_d \,:\, \widehat{\mathbf{x}} \neq \mathbf{x}} \sum_{\mathbf{y}} P(\mathbf{y} \mid \mathbf{x}) \left[\frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{s(1-1/\rho)} \left[\frac{P(\mathbf{y} \mid \widehat{\mathbf{x}})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{\lambda} \right\}^{\rho}
\end{aligned} \tag{58}$$

Consider now the GB (19), with $\mathcal{R}$ in the form

$$\mathcal{R} = \left\{ \mathbf{y} \;\middle|\; \ln \left[\frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right] \leq nR \right\} \tag{59}$$

By using the same techniques as in Section 4.2 (that is, Chernoff bounding and optimizing with respect to $\mathcal{R}$), we obtain

$$\begin{aligned}
\mathbb{P}[e_d] \;\leq\; & e^{H(\rho)} \left\{ \sum_{\mathbf{y}} \left[P(\mathbf{y} \mid \mathbf{x}) \frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{s} \right\}^{1-\rho} \\
& \times \left\{ \sum_{\widehat{\mathbf{x}} \in \mathcal{X}_d \,:\, \widehat{\mathbf{x}} \neq \mathbf{x}} \sum_{\mathbf{y}} P(\mathbf{y} \mid \mathbf{x}) \left[\frac{f(\mathbf{y})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{r} \left[\frac{P(\mathbf{y} \mid \widehat{\mathbf{x}})}{P(\mathbf{y} \mid \mathbf{x})} \right]^{s(1-1/\rho)} \right\}^{\rho}
\end{aligned} \tag{60}$$

which is (58) multiplied by a factor $\exp H(\rho)$. Thus, removing $\exp H(\rho)$ from (60) we still have an upper bound, viz., the one obtained from (58).

References

[1] E. Agrell, "Voronoi regions for binary linear block codes," *IEEE Trans. Inform. Theory*, Vol. 42, No. 1, pp. 310–316, Jan. 1996.

[2] E. Agrell, "On the Voronoi neighbor ratio for binary linear block codes," *IEEE Trans. Inform. Theory*, Vol. 44, No. 7, pp. 3064–3072, Nov. 1998.

[3] A. Ashikhmin and A. Barg, "Minimal vectors in linear codes," *IEEE Trans. Inform. Theory*, Vol. 44, No. 5, pp. 2010–2017, Sept. 1998.

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

[5] D. Divsalar, "A simple tight bound on error probability of block codes with application to turbo codes," *1999 IEEE Communication Theory Workshop*, Aptos, CA, May 23–26, 1999. Also published as JPL TMO Progress Report 42-139, Nov. 15, 1999.

[6] D. Divsalar, and E. Biglieri "Upper bounds to error probabilities of coded systems beyond the cutoff rate," *Proceedings of IEEE International Symposium on Information Theory*, 25-30 June 2000, pp. 288, Italy.

[7] S. Dolinar, L. Ekroot, and F. Pollara, "Improvements on probability of error bounds for block codes on the Gaussian channel," *Proc. of the IEEE International Symposium on Information Theory (ISIT 1994)*, Trondheim, Norway, p. 243, June 27–July 1, 1994.

[8] T. M. Duman and M. Salehi, "New performance bounds for turbo codes," *IEEE Trans. Commun.*, Vol. 46, No. 6, pp. 717–723, June 1998.

[9] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, Vol. 37, No. 5, pp. 1241–1260, Sept. 1991.

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

[11] R. G. Gallager,"A simple derivation of the coding theorem and some spplications", *IEEE Trans. Inform. Theory*, Vol. IT-12, pp. 3–18, January 1965.

[12] B. Hughes, "On the error probability of signals in additive white Gaussian noise," *IEEE Trans. Inform. Theory*, Vol. 37, No. 1, pp. 151–155, Jan. 1991.

[13] T.-Y. Hwang, "Decoding linear block codes for minimizing word error rates," *IEEE Trans. Inform. Theory*, Vol. IT-25, No. 6, pp. 733–737, Nov. 1979.

[14] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inform. Theory*, Vol. 40, No. 10, pp. 1261–1271.

[15] I. Sason and S. Shamai, "Improved upper bounds on the performance of parallel and serial concatenated turbo codes via their ensemble distance spectrum," *Proc. of the IEEE International Symposium on Information Theory (ISIT 1998)*, Cambridge, MA, p. 30, August 16–21, 1998.

[16] I. Sason and S. Shamai, "Distance spectrum based improved upper bounds for parallel and serial concatenated turbo codes," *Proc. of MELECON'98, 9th International Electrotechnical Conference*, Tel Aviv, Israel, pp. 856–860, May 18–20, 1998.

[17] I. Sason and S. Shamai (Shitz), "Improved upper bounds on the ensemble performance of ML decoded low density parity check codes", *IEEE Commun. Lett.*, Vol. 4, pp. 89–91, Mar. 2000.

[18] I. Sason and S. Shamai, "Bounds on the error probability of ML decoding for block and turbo-block codes," *Annales des Télécommunications*, Vol. 54, No.3–4, pp.183–200, March–April 1999.

[19] I. Sason and S. Shamai (Shitz), "On improved Bounds on the decoding error probabilities of block codes over interleaved fading channels, with applications to turbo-like codes", *IEEE Trans. on Information Theory*, Vol. 47, No. 6, September 2001.

[20] I. Sason and S. Shamai (Shitz)and D. Divsalar, "On simple and tight upper bounds on the ML decoding error probability for block codes over interleaved fading channels", *Proceedings 6th International Symposium on Communication Theory and Applications*, 15-20 July 2001, Ambleside, UK.

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

[22] A. M. Viterbi and A. J. Viterbi, "An improved union bound for binary input linear codes on the AWGN channel, with applications to turbo decoding," *Proc. of the IEEE Information Theory Workshop*, San Diego, CA, February, 1998.

[23] A. J. Viterbi and A. M. Viterbi, "New results on serial concatenated and accumulated-convolutional turbo code performance," *Annales des Télécommunications*, Vol. 54, No. 3–4, pp. 173–182, March–April 1999.

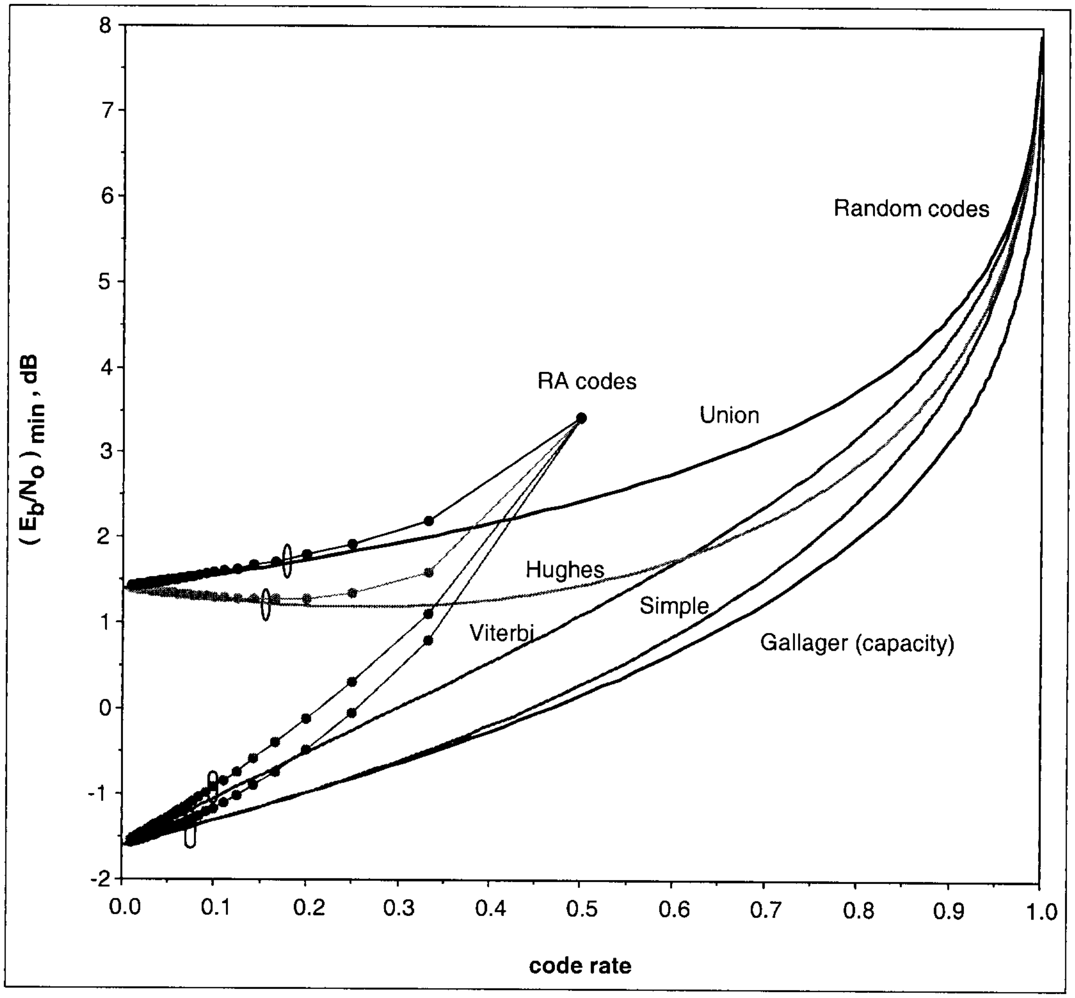

Figure 1: *Comparison of different bounds: minimum signal-to-noise ratio necessary to achieve a vanishingly small error probability over the AWGN channel for different code rates as the code block length grows to infinity. Random codes and RA codes are used. "SIMPLE:" bound described in this paper. "VITERBI:" see [22]. "HUGHES:" see [12]. "Gallager (capacity):" see [11].*

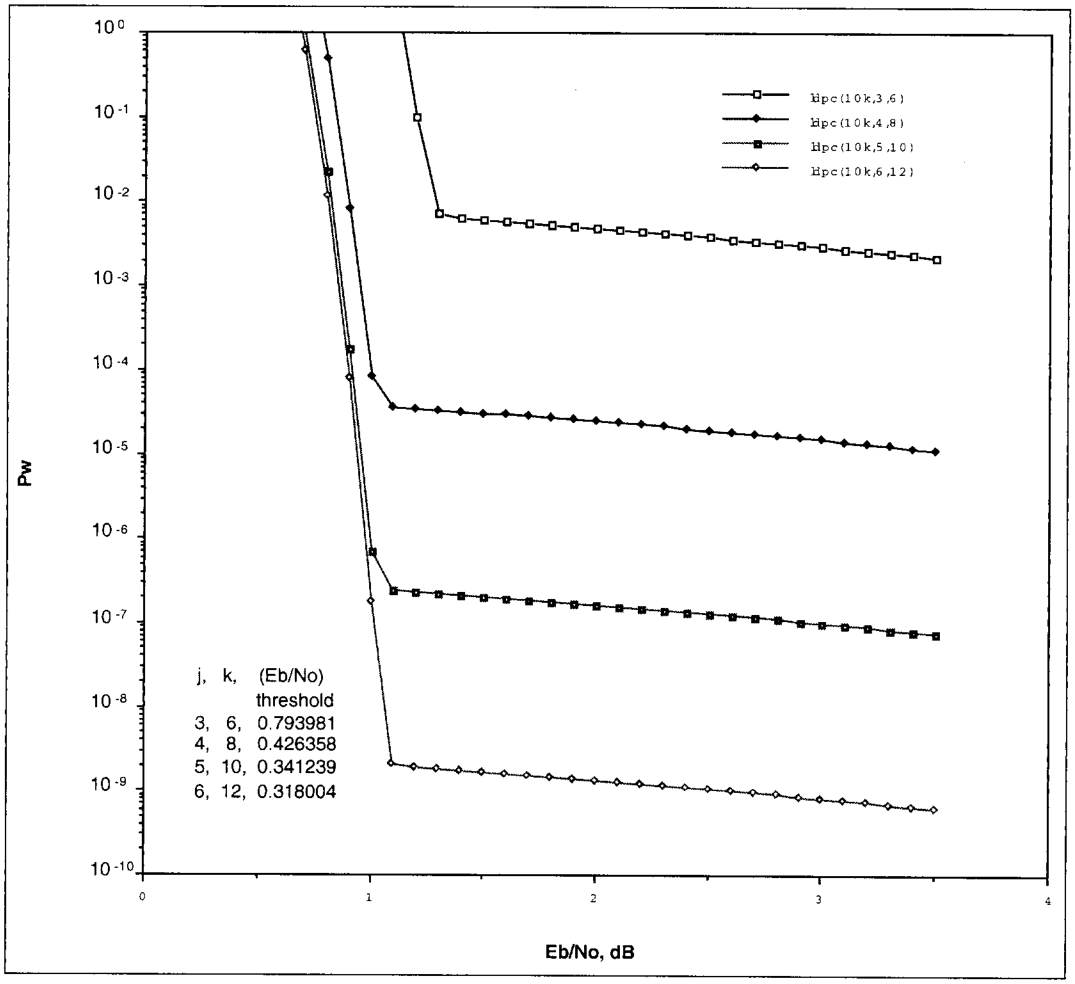

Figure 2: *Performance of LDPC codes over the AWGN channel: simple closed-form bound.*

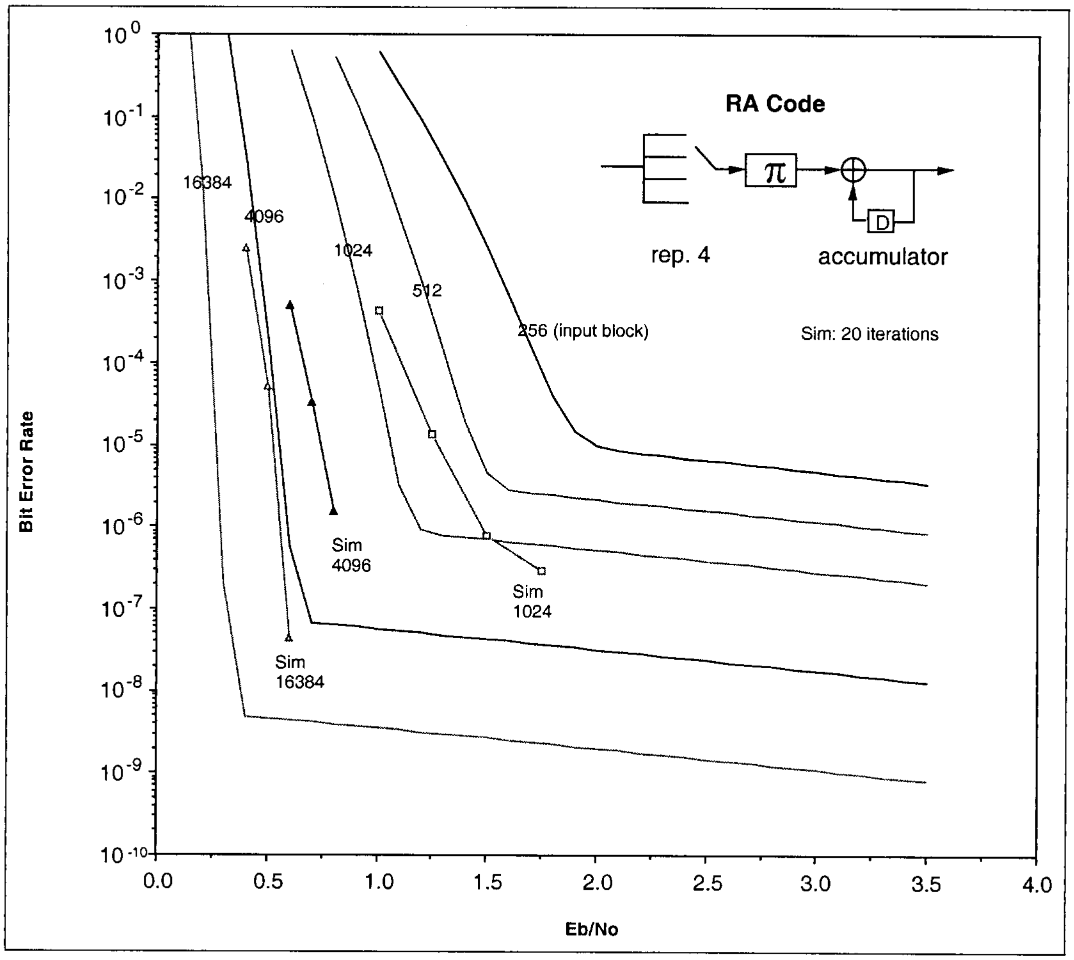

Figure 3: *Performance of rate-1/4 RA codes over the AWGN channel: simple closed-form bound for ML decoding, and simulated values for the iterative turbo decoder.*

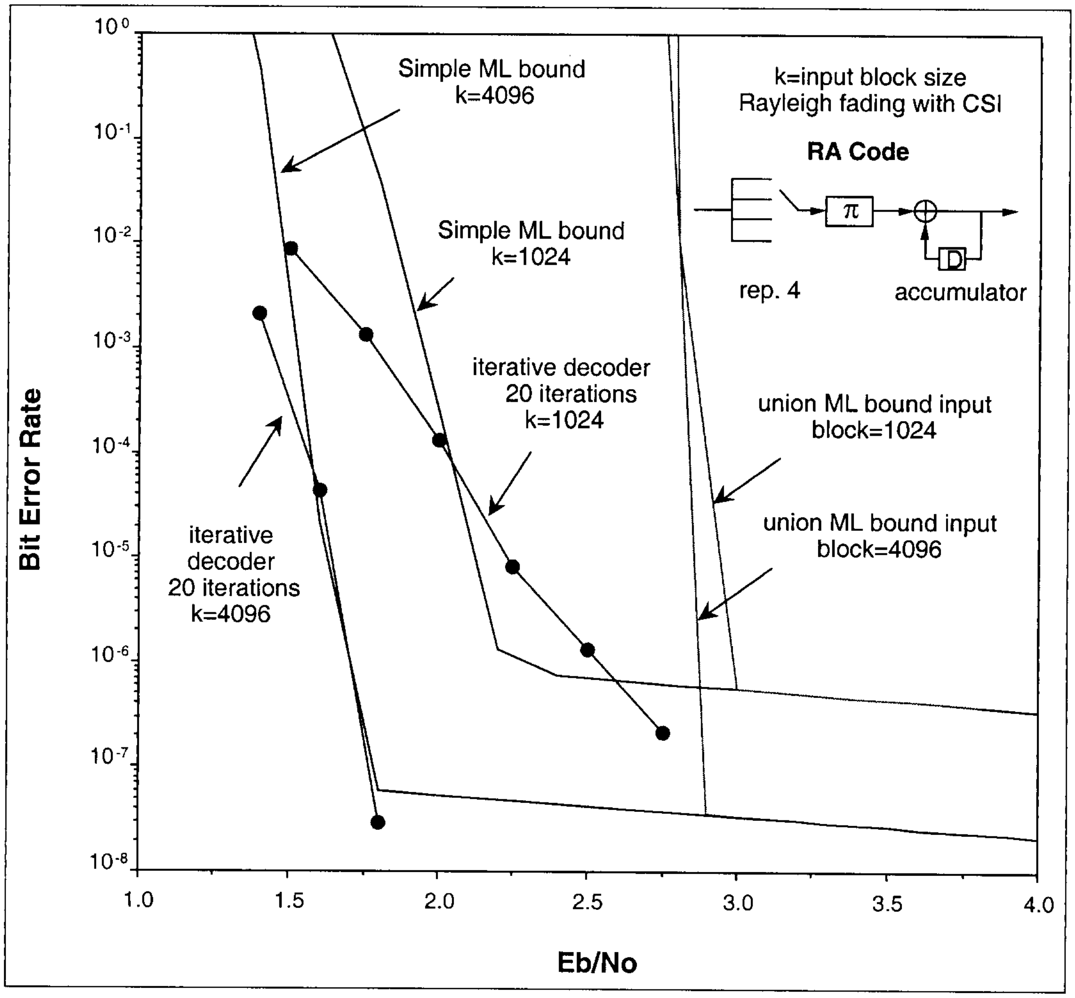

Figure 4: *Performance of rate-1/4 RA codes over the Rayleigh fading channel with perfect Channel State Information (CSI): simple closed-form bound for ML decoding, and simulated values for the iterative turbo decoder.*