A Simple Tight Bound on Error Probability of Block Codes with Application to Turbo Codes

D. Divsalar[1]

A simple bound on the probability of decoding error for block codes is derived in closed form. This bound is based on the bounding techniques developed by Gallager. We obtained an upper bound both on the word-error probability and the bit-error probability of block codes. The bound is simple, since it does not require any integration or optimization in its final version. The bound is tight since it works for signal-to-noise ratios (SNRs) very close to the Shannon capacity limit. The bound uses only the weight distribution of the code. The bound for nonrandom codes is tighter than the original Gallager bound and its new versions derived by Sason and Shamai and by Viterbi and Viterbi. It also is tighter than the recent simpler bound by Viterbi and Viterbi and simpler than the bound by Duman and Salehi, which requires two-parameter optimization. For long blocks, it competes well with more complex bounds that involve integration and parameter optimization, such as the tangential sphere bound by Poltyrev, elaborated by Sason and Shamai, and investigated by Viterbi and Viterbi, and the geometry bound by Dolinar, Ekroot, and Pollara. We also obtained a closed-form expression for the minimum SNR threshold that can serve as a tight upper bound on maximum-likelihood capacity of nonrandom codes. We also have shown that this minimum SNR threshold of our bound is the same as for the tangential sphere bound of Poltyrev. We applied this simple bound to turbo-like codes.

I. Introduction

Turbo codes proposed by Berrou et al. represent a recent breakthrough in coding theory [19], which has stimulated a large amount of new research. These codes are parallel concatenated convolutional codes (PCCC) whose encoder is formed by two [22] or more [20] *constituent* systematic encoders joined through one (or more) interleavers. Other types of turbo-like code concatenation with interleavers, such as serial concatenation of two [23] and three codes [25], hybrid concatenated codes [21], self-concatenated codes [24], and repeat accumulate (RA) codes [7] were proposed. These codes were analyzed by using the union bound, which cannot predict the performance above the cutoff rate. The performance of

[1] Communications Systems and Research Section.

these concatenated codes is close to Shannon's capacity limit for moderate to large block sizes, so there is a great demand to have bounds on performance that are useful for rates above the cutoff rate. A tight bound by Poltyrev [15] and the geometry bound proposed by Dolinar, Ekroot, and Pollara are among the tightest bounds. Recently, Duman and Salehi [14] have proposed a bound without integration that requires two-parameter optimization. Then Viterbi and Viterbi [3] developed a simple bound without a need for parameter optimization and integration. Both bounds use the second Gallager bounding technique [2]. Sason and Shamai [8–10] applied the tangential sphere bound, and Viterbi and Viterbi [5] applied a similar bound to analyze the performance of turbo codes. Both bounds are based on Poltyrev results [15]. These new tight bounds essentially use "a basic bounding technique" first developed by Gallager in 1963 [1], namely, given a transmitted codeword,

$$\Pr\{\text{word error}\} \leq \Pr\{\text{word error}, \mathbf{y} \in \Re\} + \Pr\{\mathbf{y} \notin \Re\} \tag{1}$$

where $\mathbf{y}$ is the observation vector (transmitted codeword plus noise) and $\Re$ is a region (volume) in the observation space around the transmitted codeword (this is our geometric interpretation of Gallager's basic bounding technique, which we call Gallager's first bounding technique). In [5,8,15,18], $\Re$ was defined as a cone. Gallager's $\Re$ [1] is a complicated region in observation space to be discussed in Section III.

In this article, we propose a simple upper bound on word- and bit-error probability using the basic bounding technique and the Chernov bounds proposed by Gallager. Although we used the approach and derivations of Gallager, we will show that our derived bound for nonrandom codes, in addition to its simplicity, is tighter than the original Gallager bound and versions of the Gallager bound proposed by Sason and Shamai [10] and Viterbi and Viterbi [4] since the regions in these bounds are optimum for random codes but suboptimum for nonrandom codes. Also, our bound is tighter than a bound by Viterbi and Viterbi [3] and simpler than a bound by Duman and Salehi [14]. Our proposed simple bound for short blocks may not be as tight as in [5,9,18], but it is as tight for very long blocks (as $n \to \infty$). We obtained a closed-form expression for the minimum signal-to-noise ratio (SNR) threshold above which the simple bound is useful. Our bound can predict the performance close to the capacity limit but cannot achieve the capacity limit for code rates strictly greater than 0. This was demonstrated for random codes as $n \to \infty$ using the expression for the minimum threshold of the bound. We obtained a minimum threshold for the tangential sphere bound that is identical to the minimum threshold of the simple bound. S. Dolinar[2] also independently obtained the minimum signal-to-noise ratio threshold based on the geometry bound in [18], which is identical to our minimum threshold. Thus, none of the bounds in [5,9,15,18] except the Gallager bound [2] can achieve Shannon capacity limit for random codes for code rates strictly greater than 0.

In Section II, we provide the derivation of the simple bound based on the first bounding technique by Gallager in 1963 [1]. In Section III, a modified Gallager bound is obtained as an upper bound on a bound by Gallager in 1965 (the second bounding technique) [2]. We have shown that this modified Gallager bound exactly matches the first bound of Gallager [1] except for a factor $e^{H(\rho)}$, $0 \leq \rho \leq 1$, where $H(\cdot)$ is binary entropy function. The main reason for obtaining a modified Gallager bound is as follows. In our examples, using a region $\Re$ that will be discussed, and using the first bounding technique of Gallager, we obtained a bound identical to the Viterbi and Viterbi bound [3] except for the factor mentioned above. In [3], the second bounding technique of Gallager was used. Thus, this observation motivated us to obtain a relation between the first and the second bounds by Gallager in 1963 and 1965, respectively. In Section IV, we compare the simple bound with other bounds for large block sizes. In this section, we show that asymptotically the simple bound is as tight as the tangential sphere bound of Poltyrev. A summary of results is presented in Section V. Examples are given in Section VI.

[2] S. Dolinar, Personal communication, Communications Systems and Research Section, Jet Propulsion Laboratory, Pasadena, California, 1999.

II. Derivation of the Simple Bound

A. A Simple Bound on P_w

Consider a linear binary (n, k) block code C with code rate $R_c = k/n$. We view concatenated codes with interleavers as block codes. Let the codewords be $\mathbf{x}_i \in C,\ i = 0, 1, 2, \cdots, 2^k - 1$. Assume that an arbitrary codeword $\mathbf{x}_0$ is the transmitted codeword over the additive white Gaussian noise (AWGN) channel. Divide the codewords $\{\mathbf{x}_i\}$ with Hamming distance h from $\mathbf{x}_0$ into subsets $\chi_h,\ h = 0, 1, 2, \cdots, n$. The cardinalities of these sets are $|\chi_h| = A_h$, where A_h is the number of codewords at distance h from $\mathbf{x}_0$. The goal is to obtain a simple upper bound on the word and then on the bit-error probability using maximum-likelihood decoding over the binary-input memoryless AWGN channel. The channel can be modeled as

$$y_j = x_{i,j}\gamma + n_j, \quad j = 1, 2, \cdots, n \tag{2}$$

where y_j is the observation sample; $x_{i,j} \in \{+1, -1\}$ is a component of transmitted codeword $\mathbf{x}_i$; n_j is a zero-mean unit variance Gaussian noise sample, $\gamma^2 = 2R_c(E_b/N_0)$; E_b is the information bit energy; and $N_0/2$ is the two-sided power spectral density of a white Gaussian noise process at the receiver. The likelihood function (channel transition probability) is

$$p(\mathbf{y}|\mathbf{x}_i) = \prod_{j=1}^{n} p(y_j|x_{i,j}), \quad i = 0, 1, 2, \cdots, 2^k - 1 \tag{3}$$

where

$$p(y_j|x_{i,j}) = \frac{1}{\sqrt{2\pi}} e^{[-(y - x_{i,j}\gamma)^2]/2}, \quad -\infty < y_j < \infty \tag{4}$$

Using the maximum-likelihood decision rule, given that $\mathbf{x}_0$ is transmitted, a decoding error occurs if

$$\sum_{j=1}^{n} y_j x_{i,j} > \sum_{j=1}^{n} y_j x_{0,j} \tag{5}$$

for some $i \neq 0$. Define an error event:

$$E_h \triangleq \text{ some } \mathbf{x} \in \chi_h \text{ is chosen in preference to the } \mathbf{x}_0 \text{ codeword} \tag{6}$$

Then, using the union bound, we have

$$P_{\text{word}} \leq \sum_{h>0} \Pr\{E_h|\ \mathbf{x}_0\} \tag{7}$$

In the following, we denote $\Pr\{E_h|\ \mathbf{x}_0\}$ by $p(h)$.

Next define an arbitrary n-dimensional region (volume) $\Re$ in the n-dimensional observation space containing the transmitted codeword, so that

$$p(h) = \Pr\{E_h, \mathbf{y} \in \Re \mid \mathbf{x}_0\} + \Pr\{E_h, \mathbf{y} \notin \Re \mid \mathbf{x}_0\} \tag{8}$$

Next we use the basic bounding technique on $p(h)$, namely,

$$p(h) \leq \Pr\{E_h, \mathbf{y} \in \Re \mid \mathbf{x}_0\} + \Pr\{\mathbf{y} \notin \Re \mid \mathbf{x}_0\} \tag{9}$$

In order to obtain a closed-form solution, we have chosen $\Re$ as an n-dimensional sphere with radius $\sqrt{n\ R^2}$ to be optimized. In order to obtain a tight bound, we choose the center of the sphere at $\eta\gamma\mathbf{x}_0$, a point along the line connecting the origin to the codeword $\mathbf{x}_0$. The translation factor η is a parameter to be optimized (see Fig. 1).

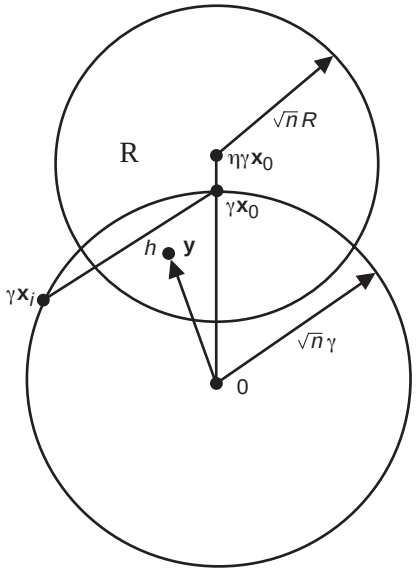

Fig. 1. Geometric interpretation of the simple bound.

The n-dimensional region (volume) then is defined as

$$\Re = \left\{\mathbf{y} \mid \ \|\mathbf{y} - \eta\gamma\mathbf{x}_0\|^2 \leq nR^2\right\} \tag{10}$$

where $\|\mathbf{y} - \eta\gamma\mathbf{x}_0\|^2 = \sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2$. Using the above-defined region and the union bound on the first term of Eq. (9), we get

$$p(h) \leq \sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \Pr\left\{\sum_{j=1}^{n} y_j x_{i,j} > \sum_{j=1}^{n} y_j x_{0,j}, \sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2 \leq nR^2 \mid \mathbf{x}_0\right\}$$

$$+ \Pr\left\{\sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2 \geq nR^2 \mid \mathbf{x}_0\right\} \tag{11}$$

At this point, we further upper bound the two terms on the right-hand side of the above equation, using the Chernov bounds described in [1], namely, for two random variables, Z and W, we have

$$\Pr\{Z \geq 0, W \leq 0\} \leq E\left\{e^{t\ Z+r\ W}\right\}, \quad t \geq 0, \quad r \leq 0 \tag{12}$$

and

$$\Pr\{W \geq 0\} \leq E\left\{e^{s\ W}\right\}, \quad s \geq 0 \tag{13}$$

where E represents expectation. Let $Z = \sum_{j=1}^{n} y_j x_{i,j} - \sum_{j=1}^{n} y_j x_{0,j}$ and $W = \sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2 - nR^2$. Using the Chernov bound, Eq. (12), with optimized parameter $t = (1-2r\eta)\gamma/2$, we get

$$\sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \Pr\left\{\sum_{j=1}^{n} y_j x_{i,j} > \sum_{j=1}^{n} y_j x_{0,j}, \sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2 \leq nR^2 \mid \mathbf{x}_0\right\}$$

$$\leq A_h e^{-nrR^2} f^h(\gamma, r, \eta) g^{n-h}(\gamma, r, \eta) \triangleq e^{-nrR^2} A \tag{14}$$

where

$$f(\gamma, r, \eta) = \frac{e^{-\frac{\gamma^2}{2}\left(1-2r\eta^2\right)}}{\sqrt{1-2r}} \tag{15}$$

and

$$g(\gamma, r, \eta) = \frac{e^{\frac{\gamma^2}{2}\frac{2r(1-\eta)^2}{1-2r}}}{\sqrt{1-2r}} \tag{16}$$

Note that $A_h = |\chi_h|$ was defined as the number of codewords at distance h from $\mathbf{x}_0$. At this point, we can average this upper bound over all possible transmitted codewords, $\mathbf{x}_0$, and use the average of A_h if the code is a nonlinear binary code. For turbo codes or turbo-like codes, we can average A_h over all possible interleavers (using the uniform interleaver concept [22]). In both cases, we still use the notation A_h for such averaged code distribution. Also, for $1 - 2s > 0$, we obtain

$$\Pr\left\{\sum_{j=1}^{n}(y_j - \eta x_{0,j}\gamma)^2 \geq nR^2 \mid \mathbf{x}_0\right\} \leq e^{-nsR^2} g^n(\gamma, s, \eta) \triangleq e^{-nsR^2} B \tag{17}$$

Minimizing the upper bound on $p(h)$ with respect to e^{nR^2}, we get

$$p(h) \leq e^{H(s/(s-r))}\ A^{s/(s-r)} B^{-r/(s-r)} \tag{18}$$

where $H(x)$ is the binary entropy function, $H(x) = -x\ln x - (1-x)\ln(1-x)$.

At this point, we redefine the parameters to be optimized. First let

$$\rho \triangleq \frac{s}{s-r} \tag{19}$$

where $0 \leq \rho \leq 1$. Then,

$$s = -r\frac{\rho}{1-\rho} \tag{20}$$

Now the upper bound will be in terms of ρ, r, and η. Next define

$$\beta \triangleq \rho(1-2r) \tag{21}$$

Since $s < 1/2$, then $0 \leq \beta \leq 1$, and define

$$\zeta \triangleq \rho(1-2r\eta) \tag{22}$$

With this redefinition of the parameters, the upper bound will be in terms of ζ, ρ, and β. Now define $\delta = h/n$, $r(\delta) \triangleq (\ln A_h)/n$, and $c = \gamma^2/2$ (note that $c = R_c E_b/N_0$). With a simple manipulation, we obtain

$$p(h) \leq e^{H(\rho)} e^{-nE(c,h,\beta,\rho,\zeta)} \tag{23}$$

where

$$E(c,h,\rho,\beta,\zeta) = -\rho r(\delta) - \frac{\rho}{2}\ln\frac{\rho}{\beta} - \frac{1-\rho}{2}\ln\frac{1-\rho}{1-\beta} + c\left[1-(1-\delta)\frac{\zeta^2}{\beta} - \frac{(1-\zeta)^2}{1-\beta}\right] \tag{24}$$

Minimizing the bound with respect to ζ, we obtain the optimum value for ζ as

$$\zeta^* = \frac{\beta}{\beta + (1-\delta)(1-\beta)} \tag{25a}$$

and

$$\frac{\partial^2 E(c,h,\rho,\beta,\zeta)}{\partial \zeta^2} \leq 0 \tag{25b}$$

The exponent $E(c,h,\rho,\beta,\zeta^*)$ reduces to

$$E(c,h,\rho,\beta) = -\rho r(\delta) - \frac{\rho}{2}\ln\frac{\rho}{\beta} - \frac{1-\rho}{2}\ln\frac{1-\rho}{1-\beta} + \frac{\delta\beta}{1-\delta(1-\beta)}c \tag{26}$$

The optimum ρ that maximizes the exponent $E(c,h,\rho,\beta)$ is

$$\rho^* = \frac{1}{1+\frac{1-\beta}{\beta}e^{2r(\delta)}} \tag{27a}$$

and

$$\frac{\partial^2 E(c,h,\rho,\beta)}{\partial \rho^2} = -\frac{1}{2\rho(1-\rho)} \leq 0 \tag{27b}$$

Using the optimum value of ρ, we obtain

$$E(c,h,\beta) = -r(\delta) + \frac{1}{2}\ln\left[\beta + (1-\beta)e^{2r(\delta)}\right] + \frac{\delta\beta}{1-\delta(1-\beta)}c \tag{28}$$

The optimal β is the solution to the following equation:

$$\frac{\partial E(c,h,\beta)}{\partial \beta} = \frac{(1-e^{2r(\delta)})}{\beta + (1-\beta)e^{2r(\delta)}} + \frac{2\delta(1-\delta)}{[1-\delta(1-\beta)]^2}c = 0 \tag{29}$$

which results in

$$\beta^* = \sqrt{c\frac{1-\delta}{\delta}\frac{2}{1-e^{-2r(\delta)}} + \left(\frac{1-\delta}{\delta}\right)^2[(1+c)^2-1]} - \frac{1-\delta}{\delta}(1+c) \tag{30}$$

Also note that

$$\frac{\partial^2 E(c,h,\beta)}{\partial \beta^2} = -\frac{\left(1-e^{2r(\delta)}\right)^2}{(\beta+(1-\beta)e^{2r(\delta)})^2} - \frac{4\delta^2(1-\delta)}{[1-\delta(1-\beta)]^3}c \leq 0 \tag{31}$$

Using the optimum value of β in $E(c,h,\beta)$ yields the exponent $E(c,h) = E(c,h,\beta^*)$. For $\beta = 1$, we get the union bound. We are interested in obtaining the minimum signal-to-noise ratio for which the exponent of the bound is positive for all h. This may be called the minimum threshold of the bound. Thus, for fixed h, first we should have $\min_c E(c,h) \geq 0$. We see that this occurs if $\beta = 0$. Also note that $([\partial E(c,h,\beta)]/[\partial\beta])|_{\beta=0} > 0$ if

$$c > \left(1-e^{-2r(\delta)}\right)\frac{1-\delta}{2\delta} \triangleq c_0(\delta) \tag{32}$$

This implies that

$$c \geq \max_{\delta}\left(1-e^{-2r(\delta)}\right)\frac{1-\delta}{2\delta} \triangleq c_0 \text{ (Simple)} \tag{33}$$

The c_0 represents the minimum E_s/N_0 threshold and can serve as a closed-form tight upper bound on the so called "maximum-likelihood (ML) capacity" for families of nonrandom codes.

Note that β also can be expressed as

$$\beta^* = \frac{1-\delta}{\delta}\left[\sqrt{\frac{c}{c_0(\delta)} + (1+c)^2 - 1} - (1+c)\right] \tag{34}$$

Finally, we have

$$P_w \leq \sum_h \kappa_h e^{-nE(c,h,\beta^*)} \tag{35}$$

where

$$\kappa_h = e^{H(\beta^*/\beta^* + (1-\beta^*)e^{2r(\delta)})} \tag{36}$$

Shortly we will present a slight improvement to this bound.

For $\beta = 1$, we have the union bound, i.e.,

$$E(c,h,1) = -r(\delta) + \delta c \tag{37}$$

where

$$c \geq \max_\delta \frac{r(\delta)}{\delta} \triangleq c_0(\text{union}) \tag{38}$$

In this case, the $c_0(\text{union})$ represents the cutoff-rate threshold for random codes. For random codes, $r(\delta) = H(\delta) - (1 - R_c)\ln 2$; thus, $c_0(\text{union}) = R_c(E_b/N_0)_{\min} = -\ln[2^{1-R_c} - 1]$.

At this point, it is interesting to note that, for $\mathbf{x} \in \chi_h$, the optimum translation factor is

$$\eta^* = \frac{1 + \dfrac{1}{2c_0(\delta)}}{1 + \dfrac{\delta}{1-\delta}\beta} \tag{39}$$

This means that the center of the sphere is located at $\eta^*\gamma\mathbf{x}_0$ for the set of codewords at Hamming distance h from $\mathbf{x}_0$.

B. A Simple Bound on P_b

In this subsection, we derive a simple bound on the bit-error probability P_b that uses input–output weight distributions. For an $(n,\ k)$ block code C, consider an encoder that maps each k-bit information block into an n-bit codeword $\mathbf{x} \in C$. For each codeword $\mathbf{x_i} \in C$, denote this information block by $\mathbf{u}_i$ and its mth component by $u_{i,m},\ m = 1, 2, \cdots, k$. Define an error event:

$$E_{h,m} \triangleq \{\text{ some } \mathbf{x}_i \in \chi_h \text{ is chosen in preference to } \mathbf{x}_0 \text{ codeword, and } u_{i,m} \neq u_{0,m}\} \tag{40}$$

Then, using the union bound, we have

$$P_b \leq \frac{1}{k}\sum_{m=1}^{k}\sum_h \Pr\{E_{h,m}|\ \mathbf{x}_0\} \tag{41}$$

Consider an arbitrary n-dimensional region (volume) $\Re$ in the n-dimensional observation space containing the transmitted codeword. Using the basic bounding technique, we have

$$\Pr\{E_{h,m}|\ \mathbf{x}_0\} \leq \Pr\{E_{h,m}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} + \Pr\{\mathbf{y} \notin \Re \mid \mathbf{x}_0\} \tag{42}$$

Define a pairwise error event:

$$E_{\mathbf{x}_i,\mathbf{x}_0} \triangleq \{\ \mathbf{x}_i \in \chi_h \ \text{ is chosen in preference to } \mathbf{x}_0\} \tag{43}$$

and

$$\begin{aligned}\Pr\{E_{h,m}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} &\leq \sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \Pr\{u_{i,m} \neq u_{0,m},\ E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} \\ &= \sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \Pr\{u_{i,m} \neq u_{0,m} \mid E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re, \mathbf{x}_0\} \\ &\quad \times \Pr\{E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\}\end{aligned} \tag{44}$$

Define an indicator function:

$$I_{u_{i,m},u_{0,m}} = \begin{cases} 1 & \text{if } u_{i,m} \neq u_{0,m} \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

Then,

$$\Pr\{u_{i,m} \neq u_{0,m} \mid E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re, \mathbf{x}_0\} = I_{u_{i,m},u_{0,m}} \tag{46}$$

Thus,

$$\frac{1}{k}\sum_{m=1}^{k} \Pr\{E_{h,m}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} \leq \sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \sum_{m=1}^{k} \frac{I_{u_{i,m},u_{0,m}}}{k} \Pr\{E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} \tag{47}$$

If $\Re$ is sufficiently symmetric, then $\Pr\{E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\}$ does not depend on the particular $\mathbf{x}_i \in \chi_h$, so

$$\frac{1}{k}\sum_{m=1}^{k} \Pr\{E_{h,m}|\ \mathbf{x}_0\} \leq \sum_{w=1}^{k} \frac{w}{k} A_{w,h} \Pr\{E_{\mathbf{x}',\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} + \Pr\{\mathbf{y} \notin \Re \mid \mathbf{x}_0\} \tag{48}$$

for any $\mathbf{x}' \in \chi_h$. The input–output coefficients $A_{w,h}$ are defined as the number of codewords, $\mathbf{x}_i$, at distance h from $\mathbf{x}_0$, with $d_H(\mathbf{u}_i, \mathbf{u}_0) = w$, where $d_H(\cdot,\cdot)$ represents the Hamming distance between two binary sequences. For turbo or turbo-like codes, $A_{w,h}$ is averaged over all interleavers. Define

$$r_b(\delta) = \frac{\ln\left\{\sum_{w=1}^{k} \frac{w}{k} A_{w,h}\right\}}{n} \tag{49}$$

Then the upper bound derived for P_w can be used for P_b if we replace $r(\delta)$ with $r_b(\delta)$. In the above derivations, we have not restricted ourselves to any particular region $\Re$, so this technique can be applied to other types of regions as well.

If we choose

$$\Re = \{\mathbf{y} \mid \ \|\mathbf{y} - \eta\gamma\mathbf{x}_0\|^2 \leq nR^2\} \tag{50}$$

and use the results on P_w from the previous subsection, we obtain

$$P_b \leq \sum_h \kappa_{b,h} e^{-nE_b(c,\delta,\beta)} \tag{51}$$

where

$$E_b(c,\delta,\beta) = -r_b(\delta) + \frac{1}{2}\ln\left[\beta + (1-\beta)e^{2r_b(\delta)}\right] + \frac{\delta\beta}{1-\delta(1-\beta)}c \tag{52}$$

where

$$\beta = \sqrt{c\frac{1-\delta}{\delta}\frac{2}{1-e^{-2r_b(\delta)}} + \left(\frac{1-\delta}{\delta}\right)^2[(1+c)^2-1]} - \frac{1-\delta}{\delta}(1+c) \tag{53}$$

and

$$\kappa_{b,h} = e^{H\left(\beta/[\beta+(1-\beta)e^{2r_b(\delta)}]\right)} \tag{54}$$

For $\beta = 1$, the simple bound reduces to the union bound.

C. Summary of the Simple Bound

The derived bound both for word- and bit-error probabilities can be further tightened. The coefficient factors κ_h and $\kappa_{b,h}$ can be ignored (i.e., set to 1), as will be discussed shortly. Furthermore, it is shown in [27] that, for improved union-type bounds, the range of h can be reduced without violating the upper bound. Thus, we obtain

$$P_w \leq \sum_{h=h_{min}}^{n-k+1} \min\left\{e^{-n\ E(c,h)},\quad A_h Q\left(\sqrt{2ch}\right)\right\} \tag{55}$$

and

$$P_b \leq \sum_{h=h_{min}}^{n-k+1} \min\left\{e^{-n\ E_b(c,\delta)},\quad \sum_w \frac{w}{k} A_{w,h} Q\left(\sqrt{2ch}\right)\right\} \tag{56}$$

III. On Gallager Bounds

In [13], Sason and Shamai made several observations and extensions on Gallager's 1963 bound [1]. Although versions of the Gallager bound with one or even two parameters are asymptotically tight for random codes (they achieve the capacity limits as $n \to \infty$), they are not tight for nonrandom codes. The reason is that the region $\Re$ used in the bound is optimum for random codes, but it is not optimum for nonrandom codes. In his thesis, Gallager indeed introduced a function of observation sample (he denoted it by $f(y)$) to be optimized to obtain a tight bound. We will shortly see that this function is related to the region $\Re$. So, optimizing this function is equivalent to optimizing the region $\Re$. Gallager found the optimal $f(y)$ using calculus of variations. The result for optimum $f(y)$ was not in a closed-form expression. However, the result reduces to a closed-form solution for random codes with two parameters for optimization.

A. Modified Gallager Bound

In 1965, Gallager [2] proposed a second bounding technique on word-error probability, given by

$$P_w \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{57}$$

where the outer sum is over the space of channel output vectors and would be replaced by an integral when the output of the channel is continuous. This bound was modified by Duman and Salehi [14] as follows: Define a nonnegative function $f(\mathbf{y})$. Then,

$$P_w \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{58}$$

$$= \sum_{\mathbf{y}} f(\mathbf{y}) \left\{ \sum_{m' \neq m} \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{-(1/\rho)} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{59}$$

If $f(\mathbf{y})$ represents a density function, i.e., $\sum_{\mathbf{y}} f(\mathbf{y}) d\mathbf{y} = 1$, then for parameter $0 \leq \rho \leq 1$, the above bound, using Jensen's inequality, can be further upper bounded as

$$P_w \leq \left\{ \sum_{m' \neq m} \sum_{\mathbf{y}} f(\mathbf{y}) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{-(1/\rho)} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{60}$$

$$= \left\{ \sum_{m' \neq m} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{1-(1/\rho)} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{61}$$

Sason and Shamai [11], using calculus of variation and an iterative method similar to that proposed by Gallager [1], developed an optimum $f(\mathbf{y})$ without symmetry. In the next subsection, we further modify the above bound.

B. Relation Between the First and the Second Bounding Techniques of Gallager

Based on the results of Duman and Salehi [14],the second bounding technique of Gallager [2] can be further modified by introducing another parameter, $s \geq 0$, as follows:

$$P_w \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho}$$

$$\leq \left\{ \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{s} \right\}^{(1-\rho)}$$

$$\times \left\{ \sum_{m' \neq m} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{s(1-[1/\rho])} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{62}$$

This can be shown by replacing $f(\mathbf{y})$ in Eq. (59) with $f^s(\mathbf{y})/P^{s-1}(\mathbf{y}|\mathbf{x}_m)$ and then using the argument of density function and Jensen's inequality. Since $f(\mathbf{y})$ multiplied by a positive constant does not change the bound, we can choose $f(\mathbf{y})$ such that

$$\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{s} = 1 \tag{63}$$

This allows us to express the bound in Eq. (62) alternatively as

$$P_w \leq \left\{ \sum_{m' \neq m} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{s(1-[1/\rho])} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{64}$$

Next we consider the first bounding technique of Gallager. In his first bounding technique [1], the region $\Re$ can be defined as

$$\Re = \left\{ \mathbf{y} \mid \ln \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right] \leq nR \right\} \tag{65}$$

This provides a geometric interpretation for the Gallager bound. Then the word-error probability can be bounded as

$$P_w \leq \Pr\{\text{word error}, \mathbf{y} \in \Re\} + \Pr\{\mathbf{y} \notin \Re\} \tag{66}$$

The two terms on the right-hand side of the above equation, using the Chernov bounds as previously described, can be bounded as

$$\Pr\{\text{word error}, \mathbf{y} \in \Re\} \leq \sum_{m' \neq m} E\left\{ e^{t\ Z_{m'} + r\ W} \right\}, \quad t \geq 0, \quad r \leq 0 \tag{67}$$

and

$$\Pr\{\mathbf{y} \notin \Re\} \leq E\left\{e^{s\ W}\right\}, \quad s \geq 0 \tag{68}$$

where

$$Z_{m'} = \ln \frac{P\left(\mathbf{y}|\mathbf{x}_{m'}\right)}{P\left(\mathbf{y}|\mathbf{x}_{m}\right)} \tag{69}$$

and

$$W = \ln\left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)}\right] - nR \tag{70}$$

Minimizing the bound with respect to R, we obtain

$$P_w \leq e^{H(\rho)} \left\{\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)}\right]^s\right\}^{(1-\rho)} \tag{71}$$

$$\times \left\{\sum_{m' \neq m} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)}\right]^r \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)}\right]^t\right\}^{\rho} \tag{72}$$

If we define $r = s(1 - [1/\rho])$ and rename t by λ, then comparing this upper bound with the bound in Eq. (62), we see that the two bounds are identical except for a constant, $e^{H(\rho)}$, that is between 1 and 2. Thus, the first bound of Gallager [1] is an upper bound to his second bound [2]. Extensions of these results to improved union-type bounds will be discussed in Subsection III.D. The results of this subsection suggest that if the first bounding technique of Gallager is used to derive the upper bound on decoding error probability, then the constant $e^{H(\rho)}$ can be ignored (set to 1).

C. Bit-Error Probability for the Modified Gallager Bound

For an (n, k) block code C, consider an encoder that maps each k-bit information block into an n-bit codeword $\mathbf{x} \in C$. For each codeword $\mathbf{x}_m \in C$, denote this information block by $\mathbf{u}_m$ and its ith component by $u_{m,i}$, $i = 1, 2, \cdots, k$.

The bit-error probability can be written as

$$P_b \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \sum_{i=1}^{k} \frac{1}{k} \phi_{m,i}(\mathbf{y}) \tag{73}$$

where the indicator function, $\phi_{m,i}(\mathbf{y})$, is defined as

$$\phi_{m,i}(\mathbf{y}) = \begin{cases} 1 & \text{if } P(\mathbf{y}|\mathbf{x}_{m'}) \geq P(\mathbf{y}|\mathbf{x}_m) \text{ and } u_{m',i} \neq u_{m,i} \text{ for some } m' \neq m \\ 0 & \text{otherwise} \end{cases} \tag{74}$$

Also define an indicator function:

$$I(u_{m',i}, u_{m,i}) = \begin{cases} 1 & \text{if } u_{m',i} \neq u_{m,i} \\ 0 & \text{otherwise} \end{cases} \tag{75}$$

We now upper bound $\phi_{m,i}(\mathbf{y})$ using arguments similar to those in [2]:

$$\phi_{m,i}(\mathbf{y}) \leq \left\{ \sum_{m' \neq m} I(u_{m',i}, u_{m,i}) \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho}, \quad \rho \geq 0, \quad \lambda \geq 0 \tag{76}$$

We note that the upper bound on $\phi_{m,i}(\mathbf{y})$ is always nonnegative. Thus, when $\phi_{m,i}(\mathbf{y}) = 0$, the upper bound in Eq. (76) holds. When $\phi_{m,i}(\mathbf{y}) = 1$, for at least one m' we have

$$I(u_{m',i}, u_{m,i}) \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \geq 1 \tag{77}$$

Thus, the sum over m' is greater than or equal to 1, and raising the sum to the ρ power is still greater than one. Next, for $0 \leq \rho \leq 1$, using Jensen's inequality, we can further upper bound the average of $\phi_{m,i}(\mathbf{y})$ over i as

$$\sum_{i=1}^{k} \frac{1}{k} \phi_{m,i}(\mathbf{y}) \leq \left\{ \sum_{m' \neq m} \sum_{i=1}^{k} \frac{1}{k} I(u_{m',i}, u_{m,i}) \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{78}$$

Using this bound, we have

$$P_b \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \sum_{i=1}^{k} \frac{1}{k} I(u_{m',i}, u_{m,i}) \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho}, \quad 0 \leq \rho \leq 1, \quad \lambda \geq 0 \tag{79}$$

Denote the Hamming distance between information blocks $\mathbf{u}_{m'}$, and $\mathbf{u}_m$ as

$$d_H(\mathbf{u}_{m'}, \mathbf{u}_m) = \sum_{i=1}^{k} I(u_{m',i}, u_{m,i}) \tag{80}$$

Then,

$$P_b \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \frac{d_H(\mathbf{u}_{m'}, \mathbf{u}_m)}{k} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho}, \quad 0 \leq \rho \leq 1, \quad \lambda \geq 0 \tag{81}$$

D. Union-Type Bounds for Modified Gallager Bounds

As in Section II, let χ_h be the set of codewords with Hamming distance h from $\mathbf{x}_m$, $h = 0, 1, 2, \cdots, n$. The cardinality of these sets is $|\chi_h| = A_h$, where A_h is the number of codewords at distance h from $\mathbf{x}_m$. Starting with the upper bound by Gallager in [2], i.e.,

$$P_w \leq \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left\{ \sum_{m' \neq m} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\lambda} \right\}^{\rho} \tag{82}$$

and using the results of previous sections, the following bounds can be obtained for any nonnegative function $f(\mathbf{y})$, $0 \le \rho \le 1$, $s = -r\rho/(1-\rho) \ge 0$, $r \le 0$, and $\lambda \ge 0$:

$$P_w(\mathbf{x}_m) \le \sum_h \left[\left\{ \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^s \right\}^{(1-\rho)} \times \left\{ \sum_{\mathbf{x}_{m'} \in \chi_h : m' \ne m} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^r \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^\lambda \right\}^\rho \right] \tag{83}$$

and

$$P_b(\mathbf{x}_m) \le \sum_h \left[\left\{ \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^s \right\}^{(1-\rho)} \times \left\{ \sum_{\mathbf{x}_{m'} \in \chi_h : m' \ne m} \frac{d_H(\mathbf{u}_{m'}, \mathbf{u}_m)}{k} \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_m) \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^r \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^\lambda \right\}^\rho \right] \tag{84}$$

For binary-input discrete memoryless channels, assume $f(\mathbf{y}) = \prod_j f(y_j)$. Then we obtain

$$P_w \le \sum_h e^{-n\ E(c,h)} \tag{85}$$

where

$$E(c,h) = \max_{\rho,r,\lambda,f} \left\{ -\rho \left[r(\delta) + \delta \ln h(r,\lambda) + (1-\delta) \ln g(r) \right] - (1-\rho) \ln g\left(\frac{-r\rho}{1-\rho} \right) \right\} \tag{86}$$

and, for bit-error probability,

$$P_b \le \sum_h e^{-n\ E_b(c,h)} \tag{87}$$

where

$$E_b(c,h) = \max_{\rho,r,f} \left\{ -\rho \left[r_b(\delta) + \delta \ln h(r,\lambda) + (1-\delta) \ln g(r) \right] - (1-\rho) \ln g\left(\frac{-r\rho}{1-\rho} \right) \right\} \tag{88}$$

$r(\delta) = (\ln A_h)/n$, $r_b(\delta) = (\ln \sum_w [w/k] A_{w,h})/n$,

$$g(s) = \sum_y P\left(y|x=0\right) \left[\frac{f(y)}{P(y|x=0)} \right]^s \tag{89}$$

and

$$h(r,\lambda) = \sum_y P(y|x=0) \left[\frac{f(y)}{P(y|x=0)}\right]^r \left[\frac{P(y|x=1)}{P(y|x=0)}\right]^\lambda \tag{90}$$

We note that the exponents are similar to that obtained by Gallager in [1]; see also a report by Sason and Shamai [13]. In [1], Gallager assumed $f(y) = f(-y)$; then, for binary-input, symmetric-output channels, $\lambda = (1-r)/2$ maximizes the exponent $E(c,h)$. Further, Gallager used calculus of variation to maximize the exponent with respect to $f(y)$; by doing so, we obtain

$$\frac{f(y)}{P(y|x=0)} = k \left[\frac{1+P(y)^{1-s}}{1+2\beta P(y)^{(1-r)/2} + P(y)^{1-r}}\right]^{1/(r-s)} \tag{91}$$

where $P(y) \triangleq P(y|x=1)/P(y|x=0)$, and β should satisfy

$$\beta = \frac{\delta}{1-\delta}\,\frac{g(r,\beta)}{h(r,\beta)} \tag{92}$$

with the above-given expression for $f(y)$. The constant k is arbitrary, and its choice does not change the exponent. Let us define $\lambda = (1-r)/2 \geq 1/2$, and after using Eq. (91), let us redenote $g(s)$ by $g_1(\rho+1,\lambda,\beta)$; $g(r)$ by $g(\rho,\lambda,\beta)$; and $h(r,\lambda)$ by $h(\rho,\lambda,\beta)$.

$$g_1(\rho+1,\lambda,\beta) = \sum_y P(y|x=0) \left[\frac{1+P(y)^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta P(y)^\lambda + P(y)^{2\lambda}}\right]^{-\rho} \tag{93}$$

$$g(\rho,\lambda,\beta) = \sum_y P(y|x=0) \left[\frac{1+P(y)^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta P(y)^\lambda + P(y)^{2\lambda}}\right]^{1-\rho} \tag{94}$$

and

$$h(\rho,\lambda,\beta) = \sum_y P(y|x=0) \left[\frac{1+P(y)^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta P(y)^\lambda + P(y)^{2\lambda}}\right]^{1-\rho} P(y)^\lambda \tag{95}$$

Since we assumed $f(y) = f(-y)$ and a symmetric binary channel, then it is easy to show that

$$g_1(\rho+1,\lambda,\beta) = g(\rho,\lambda,\beta) + \beta h(\rho,\lambda,\beta) \tag{96}$$

With such an assumption, we also can use

$$g(\rho,\lambda,\beta) = \sum_{y>0} P(y|x=0) \left[\frac{1+P(y)^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta P(y)^\lambda + P(y)^{2\lambda}}\right]^{1-\rho} [1+P(y)^{2\lambda}] \tag{97}$$

and

$$h(\rho,\lambda,\beta)=\sum_{y>0}P(y|x=0)\left[\frac{1+P(y)^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta P(y)^{\lambda}+P(y)^{2\lambda}}\right]^{1-\rho}2P(y)^{\lambda} \tag{98}$$

Thus, the exponent of the bound can be expressed as

$$E(c,h)=\max_{\rho,\lambda,\beta}\left\{-\rho\left[r(\delta)+\delta\ln h(\rho,\lambda,\beta)+(1-\delta)\ln g(\rho,\lambda,\beta)\right]-(1-\rho)\ln\left[g(\rho,\lambda,\beta)+\beta h(\rho,\lambda,\beta)\right]\right\} \tag{99}$$

where $\lambda\geq 1/2$, $\delta/(1-\delta)\leq\beta\leq 1$, and $0\leq\rho\leq 1$. For random codes, it can be shown that $\beta=1$ and $\lambda=1/(1+\rho)$ minimize the bound, and the minimum SNR threshold coincides with the capacity limit. The parameter β is the solution to the nonlinear equation

$$\beta=\frac{\delta}{1-\delta}\,\frac{g(\rho,\lambda,\beta)}{h(\rho,\lambda,\beta)} \tag{100}$$

which depends on the values of ρ, λ, and δ. Rather than solving this nonlinear equation, for every ρ, λ, and δ, we treat β as a parameter to be optimized numerically to obtain the maximum exponent for each δ. Note that

$$g(\rho,\lambda,\beta)>h(\rho,\lambda,\beta) \tag{101}$$

This implies that $\beta\geq\delta/(1-\delta)$. By using an upper bound on the parameter s in Gallager's bound, which was obtained in [13], we may choose $\lambda\leq 1/2\rho$, but this requires further study.

For a binary-input AWGN channel, $P(y)=e^{-2y\sqrt{2c}}$, $P(y|x=0)=1/(\sqrt{2\pi})e^{-(1/2)(y-\sqrt{2c})^2}$, $c=E_s/N_0$, and $\sum_y$ should be replaced by $\int$. Let $z=e^{-2y\sqrt{2c}}$; then we have

$$g(\rho,\lambda,\beta)=\int_0^1\frac{e^{-\left[(\ln z+4c]^2/16c)+\ln z\right]}}{\sqrt{16c\pi}}\left[\frac{1+z^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta z^{\lambda}+z^{2\lambda}}\right]^{1-\rho}\left[1+z^{2\lambda}\right]dz \tag{102}$$

and

$$h(\rho,\lambda,\beta)=\int_0^1\frac{e^{-[(\ln z+4c]^2/16c)+\ln z]}}{\sqrt{16c\pi}}\left[\frac{1+z^{(1-2\rho\lambda)/(1-\rho)}}{1+2\beta z^{\lambda}+z^{2\lambda}}\right]^{1-\rho}2z^{\lambda}dz \tag{103}$$

For linear codes, if we do not use the symmetry $f(y)=f(-y)$, then the optimized $f(y)$ using the calculus of variation method of Gallager is obtained as

$$\frac{f(y)}{P(y|x=0)}=k[1+\beta P(y)^{\lambda}]^{1/(s-r)} \tag{104}$$

where β satisfies Eq. (100)with the above optimized $f(y)$. However, in this case,

$$g(s) = \sum_y P(y|x=0)[1+\beta P(y)^\lambda]^\rho \tag{105}$$

$$g(r) = \sum_y P(y|x=0)[1+\beta P(y)^\lambda]^{\rho-1} \tag{106}$$

and

$$h(r,\lambda) = \sum_y P(y|x=0)[1+\beta P(y)^\lambda]^{\rho-1} P(y)^\lambda \tag{107}$$

We note that $g(s)$, $g(r)$, and $h(r,\lambda)$ are independent of r. Let us define

$$g(\rho+1,\lambda,\beta) = \sum_y P(y|x=0)[1+\beta P(y)^\lambda]^\rho \tag{108}$$

and

$$h(\rho,\lambda,\beta) = \sum_y P(y|x=0)[1+\beta P(y)^\lambda]^{\rho-1} P(y)^\lambda \tag{109}$$

It is easy to show that

$$g(\rho+1,\lambda,\beta) = g(\rho,\lambda,\beta) + \beta h(\rho,\lambda,\beta) \tag{110}$$

Then,

$$E(c,h) = \max_{\rho,\lambda,\beta} \{-\rho[r(\delta) + \delta \ln h(\rho,\lambda,\beta) + (1-\delta)\ln g(\rho,\lambda,\beta)] - (1-\rho)\ln[g(\rho,\lambda,\beta) + \beta h(\rho,\lambda,\beta)]\} \tag{111}$$

Thus, when we do not use the symmetry, i.e., $f(y) = f(-y)$, our modified Gallager bound coincides with Sason and Shamai's bound [13] (called "the second version of Duman and Salehi's bound"). The minimum SNR threshold of this bound meets the capacity limit for random codes for $\beta = 1$ and $\lambda = 1/(1+\rho)$.

The minimum SNR threshold for a given code-weight distribution can be obtained as the smallest c such that

$$\min_{\delta} \max_{\rho,\lambda,\beta} \{-\rho[r(\delta) + \delta \ln h(\rho,\lambda,\beta) + (1-\delta)\ln g(\rho,\lambda,\beta)] - (1-\rho)\ln[g(\rho,\lambda,\beta) + \beta h(\rho,\lambda,\beta)]\} \geq 0 \tag{112}$$

This threshold, although more complex to calculate, is tighter than all other thresholds discussed in this article, particularly for high-rate codes.

IV. Comparisons with Other Bounds

A. Simplified Sphere Bound of Hughes

This is a special case of our simple bound when the center of the sphere is located at the transmitted codeword, i.e., $\eta = 1$. In this case, the region $\Re$ is (see Fig. 2)

$$\Re = \{\mathbf{y} \mid \ \|\mathbf{y} - \gamma\mathbf{x}_0\|^2 \le nR^2\} \tag{113}$$

By setting $\eta = 1$, or equivalently setting $\zeta = \beta$ in our results, we obtain a looser bound with exponent

$$E(c, h, \beta) = -r(\delta) + \frac{1}{2}\ln\left[\beta + (1-\beta)e^{2r(\delta)}\right] + \delta\beta c \tag{114}$$

where the optimum β is

$$\beta = \frac{2\delta c - \left(1 - e^{-2r(\delta)}\right)}{2\delta c\left(1 - e^{-2r(\delta)}\right)} \tag{115}$$

and we should have

$$c \ge \max_{\delta}\left(1 - e^{-2r(\delta)}\right)\frac{1}{2\delta} \triangleq c_0 \text{ (Hughes)} \tag{116}$$

For $\beta = 1$, we have the union bound.

This bound with the exponent given by Eq. (114) can be considered as a simplified version of a bound by Hughes [17]. It is easy to show that the exponent given by Eq. (114) is smaller than our exponent for the simple bound [see Eq. (28)]. Also, c_0 (Hughes) $\ge c_0$ (Simple).

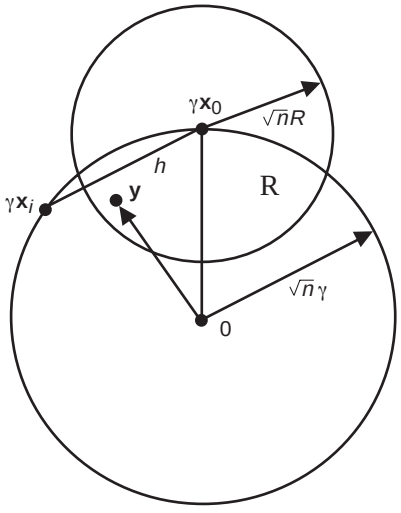

Fig. 2. Geometric interpretation of the Hughes bound.

B. Geometric Interpretation of the Viterbi-and-Viterbi Bound, and a Simplified Version of the Tangential Bound of Berlekamp

If we define the region $\Re$ as

$$\Re = \{\mathbf{y} \mid \ \langle \mathbf{y}, \gamma \mathbf{x}_0 \rangle \geq nR\} \tag{117}$$

where $\langle \cdot \rangle$ represents the inner product, we obtain the exponent of the Viterbi-and-Viterbi bound [3]. This bound also can be considered as a simplified version of the tangential bound by Berlekamp [16] (see Fig. 3).

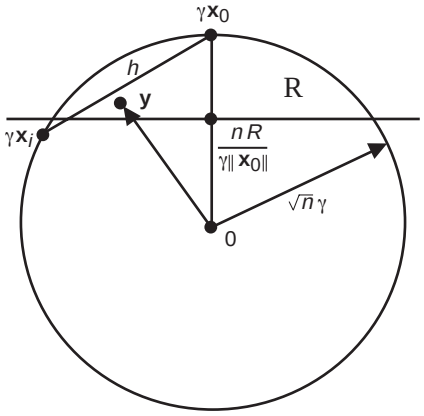

Fig. 3. Geometric interpretation of the Viterbi–Viterbi bound.

Based on our derivations, the exponent of the bound is obtained as

$$E(c, h, \rho) = -\rho r(\delta) + \frac{\delta \rho}{1 - \delta + \delta \rho} c \tag{118}$$

The optimum ρ is

$$\rho = \sqrt{\frac{(1-\delta)c}{\delta r(\delta)}} - \frac{(1-\delta)}{\delta} \tag{119}$$

Finally, for $\rho < 1$, we obtain

$$E(c, h) = \left(\sqrt{c} - \sqrt{\frac{(1-\delta)}{\delta} r(\delta)} \right)^2 \tag{120}$$

where we should have

$$c \geq \max_{\delta} r(\delta) \frac{1-\delta}{\delta} \triangleq c_0 \text{ (Viterbi-Viterbi)} \tag{121}$$

For $\rho = 1$, we have the union bound.

Note that Viterbi and Viterbi used a different approach, namely, the second bounding technique of Gallager [2], and the coefficient κ_h in their upper bound is 1. Since we got the same exponent for the bound using the first bounding technique of Gallager, this motivated us to look for a relation between the first and the second bounding techniques of Gallager, as we have shown in the previous section. Since the modified version of the Gallager bound is the same as his first bound, but without the coefficients κ_h and $\kappa_{b,h}$, these coefficients therefore can be ignored in the above-derived bounds. It is easy to show that the exponent given by Eq. (120) is small than our exponent for the simple bound. Note that as β approaches ρ in Eq. (26), we get Eq. (118). Also, c_0 (Viterbi–Viterbi) $\geq c_0$ (Simple).

C. Geometric Interpretation of the Duman–Salehi Bound, and Its Closed-Form Solution

Duman and Salehi [14] proposed the following upper bound on word-error probability (we use a " ˜ " for distinction over the parameters in Duman and Salehi's bound):

$$P_w \leq \left\{ \sum_{m' \neq m} \sum_{\mathbf{y}} F(\mathbf{y}) \left[\frac{F(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{-(1/\tilde{\rho})} \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{\tilde{\lambda}} \right\}^{\tilde{\rho}} \tag{122}$$

where

$$F(\mathbf{y}) = \prod_{i=1}^{n} \left(\frac{\tilde{\alpha}}{\pi N_0} \right)^{1/2} e^{-[(y_i - [\tilde{\beta}/\tilde{\alpha}]\sqrt{E_s}x_0)^2/(N_0/\tilde{\alpha})]} \tag{123}$$

and $\int_{-\infty}^{\infty} F(\mathbf{y}) d\mathbf{y} = 1$.

Duman and Salehi minimized the bound with respect to $\tilde{\lambda}$ and $\tilde{\beta}$ and obtained optimum solutions for these two parameters. However, the closed-form solutions for two other parameters, $\tilde{\alpha}$ and $\tilde{\rho}$, were not obtained. Thus, the bound requires numerical optimization over $\tilde{\alpha}$ and $\tilde{\rho}$ for each output weight, h, and each E_s/N_0. At first glance, it seems that our proposed bound with $s = 1$ was used; thus, it might be looser than our bound. However, since Duman and Salehi introduced two rather than one parameter in F($\mathbf{y}$), we can show that

$$\frac{F(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} = \left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)} \right]^{(1-\tilde{\alpha})/2} \tag{124}$$

Therefore, the bound can be expressed as our proposed modified Gallager bound with

$$s = \frac{1 - \tilde{\alpha}}{2} \tag{125}$$

with a new defined nonnegative function $f(\mathbf{y})$. Therefore, $\ln[f(\mathbf{y})]/[P(\mathbf{y}|\mathbf{x}_m)] \leq nR$ should define the region $\Re$. In fact,

$$\Re = \left\{ \mathbf{y} \,|\, \left\| \mathbf{y}\sqrt{\frac{2}{N_0}} \right\|^2 - 2\frac{1-\tilde{\beta}}{1-\tilde{\alpha}}\sqrt{\frac{2E_s}{N_0}} \left\langle \mathbf{y}\sqrt{\frac{2}{N_0}}, \mathbf{x}_0 \right\rangle + \zeta \leq nR \right\} \tag{126}$$

where

$$\zeta = \frac{1}{\tilde{\alpha}} \left(\frac{\tilde{\alpha} - \tilde{\beta}^2}{1 - \tilde{\alpha}} \right) \frac{2E_s}{N_0} \left|\left|\mathbf{x}_0\right|\right|^2 + \frac{0.5n}{1 - \tilde{\alpha}} \ln \tilde{\alpha} \tag{127}$$

Aside from the normalized observation and ζ, which does not depend on observation, the region $\Re$ is identical to our region. Thus, we claim that the exponent of Duman and Salehi's bound should be identical to the exponent of our bound. If this is true, then the translation factor η that we found should be equal to the translation factor in the Duman–Salehi region. Also, the parameter s in our bound should be equal to $(1 - \tilde{\alpha})/2$. Therefore, we have two equations to solve for $\tilde{\alpha}$ and $\tilde{\rho}$ (since $\tilde{\beta}$ depends on $\tilde{\rho}$), i.e.,

$$\left.\begin{aligned} s &= \frac{1 - \tilde{\alpha}}{2} \\ \eta &= \frac{1 - \tilde{\beta}}{1 - \tilde{\alpha}} \end{aligned}\right\} \tag{128}$$

If we take the first and the second derivatives of the Duman–Salehi bound with respect to $\tilde{\alpha}$ and $\tilde{\rho}$, we can in fact show that the obtained solutions for $\tilde{\alpha}$ and $\tilde{\rho}$ from the above two equation, using the results of our simple bound, satisfy the first derivative and that the second derivative is positive. Thus, these solutions minimize the bound, and the exponent of Duman and Salehi's bound will be identical to the exponent of the simple bound.

D. Simplified Version of the Tangential Sphere Bound of Poltyrev

The tangential sphere bound on word-error probability was derived by Poltyrev in [15]. The bounding technique is similar to the basic bounding technique of Gallager [1] (i.e., the first bounding technique of Gallager). For the tangential sphere bound, the region $\Re$ is a circular cone with a half angle θ whose main axis passes through the origin and the transmitted codeword (see Fig. 4).

In Fig. 4, the noise vector $\mathbf{z} = z_1, z_2, \cdots, z_n$ with the origin at $\mathbf{x}_0$ has orthogonal zero-mean unit variance components, where z_1 is along the main axis of the cone, z_2 is orthogonal to z_1, and it is in the plane

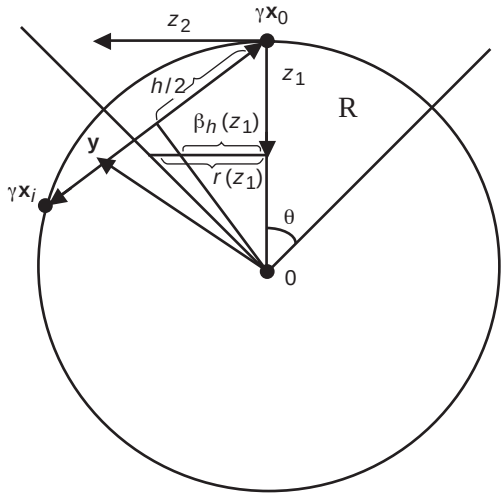

Fig. 4. Geometric interpretation of the tangential sphere bound.

defined by three points: the origin $\mathbf{0}$, the transmitted codeword $\mathbf{x}_0$, and some codeword $\mathbf{x}_i$ with Hamming distance h from the transmitted codeword. All other components of $\mathbf{z}$ are orthogonal to the described plane. Also in Fig. 4 we have

$$r(z_1) = \sqrt{\eta}\left(\sqrt{2nc} - z_1\right) \tag{129}$$

and

$$\beta_h(z_1) = \left(\sqrt{2nc} - z_1\right)\sqrt{\frac{\delta}{1-\delta}} \tag{130}$$

where $\eta = \tan^2\theta$ and $c = E_s/N_0$. In [15, Eqs. (25 and (26)], Poltyrev defines the region $\Re$ as

$$\sum_{j=2}^{n} z_i^2 \geq r^2(z_1) \tag{131}$$

over all values of z_1. (Note that this region is in fact a double cone. For a single cone, we should have the further restriction $z_1 \leq \sqrt{2nc}$, but for large n this does not make any difference). The upper bound on the probability of word error can be written as

$$P_w \leq \sum_{h} \sum_{\mathbf{x}_i \in \chi_h : i \neq 0} \Pr\{E_{\mathbf{x}_i,\mathbf{x}_0}, \mathbf{y} \in \Re \mid \mathbf{x}_0\} + \Pr\{\mathbf{y} \notin \Re \mid \mathbf{x}_0\} \tag{132}$$

If $z_2 > \beta_h(z_1)$, $E_{\mathbf{x}_i,\mathbf{x}_0}$ occurs; furthermore, the event $\{E_{\mathbf{x}_i,\mathbf{x}_0},\ \mathbf{y} \in \Re\}$ is nonempty if $z_2 \leq r(z_1)$, which implies $\beta_h(z_1) < r(z_1)$ or $\eta > \delta/(1-\delta)$. Next the upper bound is minimized over η. We can proceed by using the Chernov bounds. However, the resulting bound, although without integrations, will require numerical parameter optimization in its final version. Next we use the method of partitioning the codewords into subsets with the same Hamming distance as was used in the previous sections. Then we can have the following upper bound:

$$P_w \leq \sum_{h} \left\{ A_h \Pr\left\{ \sum_{j=2}^{n} z_i^2 \leq r^2(z_1),\ z_2 > \beta_h(z_1) \right\} + \Pr\left\{ \sum_{j=2}^{n} z_i^2 > r^2(z_1) \right\} \right\} \tag{133}$$

Now, using Chernov bounds, we get

$$\Pr\left\{ \sum_{j=2}^{n} z_i^2 > r^2(z_1) \right\} \leq \frac{\sqrt{1-2s}}{\sqrt{1+2s\eta}} e^{-nE_1(c,s,\eta)}, \quad s \geq 0 \tag{134}$$

$$E_1(c,s,\eta) = \frac{2s\eta c}{1+2s\eta} + \frac{1}{2}\ln(1-2s) \tag{135}$$

$$A_h \Pr\left\{ \sum_{j=2}^{n} z_i^2 \leq r^2(z_1),\ z_2 > \beta_h(z_1) \right\} \leq \frac{\sqrt{1-2r}}{\sqrt{1+2r\eta}} e^{-nE_2(c,r,\delta,\eta)}, \quad r \leq 0 \tag{136}$$

and

$$E_2(c,r,\delta,\eta) = c\frac{2r\eta + (1-2r)\dfrac{\delta}{1-\delta}}{1 + 2r\eta + (1-2r)\dfrac{\delta}{1-\delta}} + \frac{1}{2}\ln(1-2r) - \ r(\delta) \tag{137}$$

Rather than taking the derivative with respect to η to minimize the bound, for large n we simply solve the following equation:

$$E_1(c,s,\eta) = E_2(c,r,\delta,\eta) \tag{138}$$

The solution should result in an optimum η for very large n. Before solving the above equation, we define new parameters to simplify the expressions representing the exponents. Define

$$\rho = \frac{s}{s-r} \tag{139}$$

Hence, $0 \leq \rho \leq 1$,

$$\beta = \rho(1-2r) \tag{140}$$

$$\zeta = \frac{1}{1+2s\eta} \tag{141}$$

and

$$d = \frac{\delta}{1-\delta} \tag{142}$$

In obtaining the above bounds, we assumed $1-2s \geq 0$ and $1+2r\eta \geq 0$. This implies that $0 \leq \beta \leq 1$ and $0 \leq \zeta \leq 1$. Also, we have $\eta \geq d$. Now the exponents can be written as

$$E_1(c,\beta,\rho,\zeta) = c - c\zeta + \frac{1}{2}\ln\frac{1-\beta}{1-\rho} \tag{143}$$

and

$$E_2(c,\beta,\rho,\zeta) = c - c\,\frac{\rho\zeta}{(1+\beta d)\zeta - (1-\rho)} + \frac{1}{2}\ln\frac{\beta e^{-2r(\delta)}}{\rho} \tag{144}$$

Define $A = 1/2\ln\{[\beta(1-\rho)e^{-2r(\delta)}]/[\rho(1-\beta)]\}$. Then the solution to $E_1 = E_2$ is given as

$$\zeta = \frac{(c-(1+\beta d)A) \pm \sqrt{(c-(1+\beta d)A)^2 + 4c(1-\rho)(1+\beta d)A}}{2c(1+\beta d)} \tag{145}$$

Using the positive solution in E_1, we obtain the exponent for the simplified version of the tangential sphere bound. Next we maximize the exponent with respect to ρ. The solution to ρ is

$$\rho = \frac{\beta e^{-2r(\delta)}}{1 - \beta + \beta e^{-2r(\delta)}} \tag{146}$$

Maximizing the exponent with respect to β, we finally obtain the simplified version of the tangential sphere bound as

$$P_w \leq \sum_h \kappa e^{-nE(c,h)} \tag{147}$$

where $\kappa = (\sqrt{1-2s}/\sqrt{1+2s\eta}) + (\sqrt{1-2r}/\sqrt{1+2r\eta})$ with optimized parameters,

$$E(c,h) = \frac{1}{2}\ln(1 - \beta + \beta e^{-2r(\delta)}) + \frac{\beta d}{1+\beta d}c \tag{148}$$

and

$$\beta = \frac{1-\delta}{\delta}\left[\sqrt{\frac{c}{c_0(\delta)} + (1+c)^2 - 1} - (1+c)\right] \tag{149}$$

with

$$c_0(\delta) = \left(1 - e^{-2r(\delta)}\right)\frac{1-\delta}{2\delta} \tag{150}$$

But the exponent of the simplified tangential sphere bound is identical to the exponent of the simple bound previously derived. This implies that asymptotically the simple bound is as good as the tangential sphere bound. The minimum SNR threshold is

$$c_0 = \frac{E_s}{N_0} = \max_\delta c_0(\delta) = c_0 \text{ (Poltyrev tangential sphere)} \tag{151}$$

It is interesting to note that the optimum half angle θ is obtained as

$$\eta = \tan^2\theta = \frac{1}{2c_0(\delta)} \tag{152}$$

These results imply that, as $n \to \infty$, the optimum $\eta = \tan^2\theta$ for the tangential sphere bound is $1/2c_0$ and that $c = c_0$ represents the minimum SNR threshold for the tangential sphere bound. Moreover, the plot of c_0 for random codes monotonically diverges from the capacity plot for a binary-input AWGN channel as the code rate increases. Thus, for all code rates strictly greater than zero, the tangential sphere bound will not achieve the capacity limit for random codes. Thus, the tangential bound is loose for high code rates. All results hold for the bit-error probability if we replace $r(\delta) = \ln A_h/n$ with $r_b(\delta) = (\ln \sum_w [w/k]A_{w,h})/n$. We can also roughly show that, as $n \to \infty$, $\tan^2\theta = 1/(2c_0)$ satisfies the

following equation (the optimum θ in the tangential sphere bound satisfies this equation) that can be found, for example, in [9]:

$$\sum_{h:\eta>\delta/(1-\delta)} e^{nr(\delta)} \int_0^{\theta_\delta} \sin^{n-3}(\phi) \frac{\Gamma([n-1]/2)}{\Gamma([n-2]/2)\sqrt{\pi}} d\phi = 1 \tag{153}$$

where

$$\theta_\delta = \cos^{-1}\left(\frac{1}{\sqrt{\eta}}\sqrt{\frac{\delta}{1-\delta}}\right) \tag{154}$$

To show this, we use a result by Shannon [26], namely, as $n \to \infty$,

$$\int_0^{\theta_\delta} \sin^{n-3}(\phi) \frac{\Gamma([n-1]/2)}{\Gamma([n-2]/2)\sqrt{\pi}} d\phi \sim e^{n \ln \sin(\theta_\delta)} \tag{155}$$

Using this asymptotic result, we see that all terms in the summation will go to zero, except when the value of h or, equivalently, of δ corresponds to the maximum of $c_0(\delta)$. Furthermore, if we use the central limit theorem, then we can obtain an asymptotic expression for $\Pr\{\sum_{j=2}^n z_i^2 > r^2(z_1)\}$ and $\Pr\{\sum_{j=2}^n z_i^2 \le r^2(z_1),\ z_2 > \beta_h(z_1)\}$ as $n \to \infty$. Then one can show that, if $c \le c_0$, the tangential bound goes to 1 as $n \to \infty$. The same threshold also was obtained by S. Dolinar[3] for the $F(\theta)$ geometry bound [18] as $n \to \infty$.

1. Simplified Tangential Sphere Bound of Poltyrev Using the Modified Gallager Bound. The modified Gallager bound was derived as

$$P_w \le \sum_h \left[E\left\{\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)}\right\}^s\right]^{1-\rho} \left[\sum_{m'\neq m} E\left\{\left[\frac{f(\mathbf{y})}{P(\mathbf{y}|\mathbf{x}_m)}\right]^r \left[\frac{P(\mathbf{y}|\mathbf{x}_{m'})}{P(\mathbf{y}|\mathbf{x}_m)}\right]^\lambda\right\}\right]^\rho \tag{156}$$

Using the results of the previous subsection, we obtain this bound as

$$P_w \le \sum_h \kappa e^{-nE(c,h)} \tag{157}$$

where $\kappa = (\sqrt{1-2s}/\sqrt{1+2s\eta})^{1-\rho}(\sqrt{1-2r}/\sqrt{1+2r\eta})^\rho$ and

$$E(c,h) = \max_{\beta,\rho,\zeta}(1-\rho)\left[c - c\zeta + \frac{1}{2}\ln\frac{1-\beta}{1-\rho}\right] + \rho\left[c - c\,\frac{\rho\zeta}{(1+\beta d)\zeta-(1-\rho)} + \frac{1}{2}\ln\frac{\beta e^{-2r(\delta)}}{\rho}\right] \tag{158}$$

Maximizing over ζ, we obtain $\zeta = 1/(1+\beta d)$. The exponent reduces to

$$E(c,h) = \max_{\beta,\rho}\left[-\rho r(\delta) + \frac{\rho}{2}\ln\frac{\beta}{\rho} + \frac{1-\rho}{2}\ln\frac{1-\beta}{1-\rho} + \frac{\beta d}{1+\beta d}c\right] \tag{159}$$

[3] Ibid.

After maximizing the exponent with respect to ρ and β, we obtain the same exponent as in the previous subsection, which is identical to the exponent we obtained for the simple bound.

V. Summary of the Results on the Simple Bound

Let us express the exponent of the simple bound in another form as

$$E(c,h) = \frac{1}{2}\ln\left[1 - 2c_0(\delta)f(c,\delta)\right] + \frac{cf(c,\delta)}{1+f(c,\delta)}, \quad c_0(\delta) < c < \frac{e^{2r(\delta)}-1}{2\delta(1-\delta)} \tag{160}$$

Otherwise,

$$E(c,h) = -r(\delta) + \delta c \tag{161}$$

$\delta = h/n$, $c = R_c(E_b/N_0)$, and

$$c_0(\delta) = \left(1 - e^{-2r(\delta)}\right)\frac{1-\delta}{2\delta} \tag{162}$$

with threshold

$$c_0 = \max_{0\le\delta\le 1-R_c} c_0(\delta) \tag{163}$$

when $n \to \infty$, or

$$\left(\frac{E_b}{N_0}\right)_{min} = \frac{1}{R_c}\max_{0<\delta\le(1-R_c)} c_0(\delta) \tag{164}$$

and

$$f(c,\delta) = \sqrt{\frac{c}{c_0(\delta)} + 2c + c^2} - c - 1 \tag{165}$$

Then the upper bound on the frame-error rate (using $r(\delta) \triangleq \ln A_h/n$, where A_h is the output-weight distribution) and the upper bound on the bit-error rate (using $r(\delta) \triangleq (\ln\sum_w [w/k]A_{w,h})/n$, where $A_{w,h}$ is the input–output weight distribution) is given by

$$P_e \le \sum_{h=h_{min}}^{n-k+1} \min\left\{e^{-nE(c,h)},\ e^{nr(\delta)}Q\left(\sqrt{2ch}\right)\right\} \tag{166}$$

All other union-type bounds, including the modified Gallager bounds, can be tightened in this way, i.e., by the $Q(.)$ function and limiting the range of h up to $n-k+1$ ([27]). The simple bound and all bounds obtained by using the Chernov bounding technique can be further tightened by a technique discussed in [27]. The simple bound and the modified Gallager bound were extended to independent fading channels in [27]. The simple bound is the tightest closed-form upper bound on the decoding error rate, since c_0,

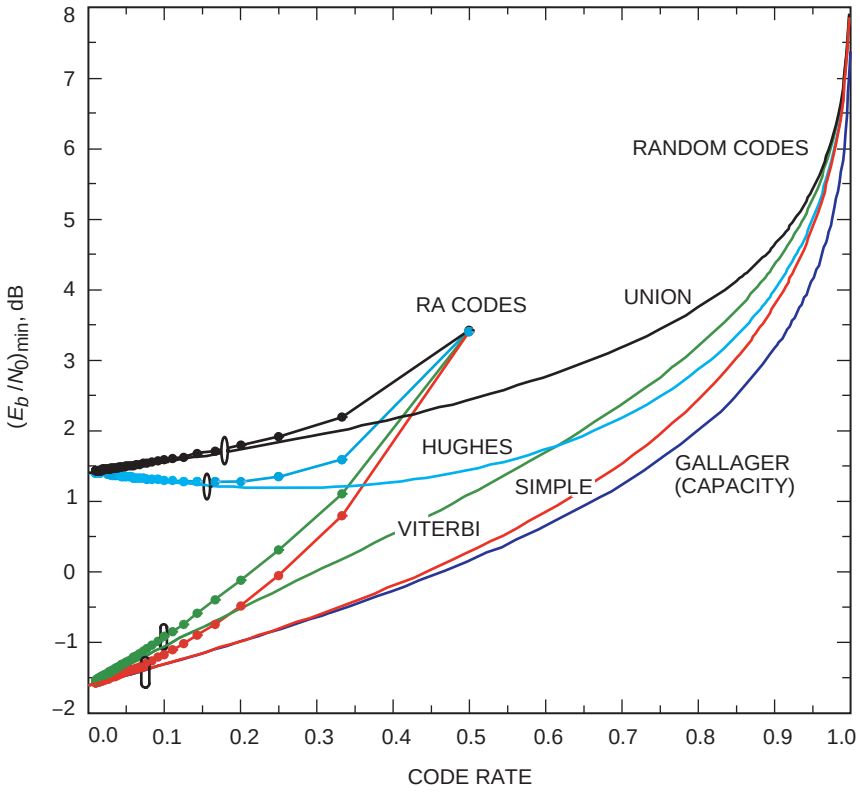

Fig. 5. Comparison of the minimum threshold, c_0, for random codes and RA codes as $n \to \infty$ for various bounds.

the minimum signal-to-noise-ratio threshold as $n \to \infty$ for various bounds, has the following relation (see also Fig. 5, where c_0's for various bounds for random codes and repeat–accumulate (RA) codes as $n \to \infty$ are compared with capacity limits):

$$\begin{aligned} c_0(\text{Simple}) &= c_0 \text{ (Poltyrev tangential sphere)} = c_0 \text{ (Dolinar–Ekroot–Pollara)} \\ &= c_0 \text{ (Duman–Salehi)} \le c_0 \text{ (Viterbi–Viterbi)} \\ &= c_0 \text{ (Berlekamp)} \le c_0 \text{ (Union)} \end{aligned}$$

Also, c_0 (Simple) $\le c_0$ (Hughes) $\le c_0$ (Union). However, c_0 (Hughes) is worse than c_0 (Viterbi–Viterbi) for low code rates, and better for high code rates where c_0's are given as

$$c_0(\text{Simple}) = \max_{\delta} \left(1 - e^{-2r(\delta)}\right) \frac{1-\delta}{2\delta} \tag{167}$$

$$c_0(\text{Viterbi-Viterbi}) = \max_{\delta} r(\delta) \frac{1-\delta}{\delta} \tag{168}$$

$$c_0(\text{Hughes}) = \max_{\delta} \left(1 - e^{-2r(\delta)}\right) \frac{1}{2\delta} \tag{169}$$

and

$$c_0(\text{Union}) = \max_{\delta} \frac{r(\delta)}{\delta} \tag{170}$$

VI. Examples

The simple bound was used to bound the ML word-error probability of rate-1/2 (n,j,k) low-density parity-check (LDPC) codes [1,12], as shown in Fig. 6. In the example, rate-1/2 (n, j, k) low-density parity-check codes for n=10,000, j=3,4,5,6, and k=2j are considered. The minimum SNR threshold using c_0 (Simple) is also shown in the figure.

In the second example, as shown in Fig. 7, the simple bound is applied to bound the ML performance of rate-1/4 repeat accumulate (RA) codes [7]. In the figure, the performance of the suboptimum iterative turbo decoder for RA codes also is shown.

VII. Conclusion

We derived a simple tight bound on bit- and frame-error rates for block codes. This is the tightest closed-form upper bound on the decoding error rate. The minimum E_b/N_0 threshold can be computed as $(n \to \infty)$:

$$\left(\frac{E_b}{N_0}\right)_{\text{threshold}} = \max_{\delta} \frac{c_0(\delta)}{R_c} \tag{171}$$

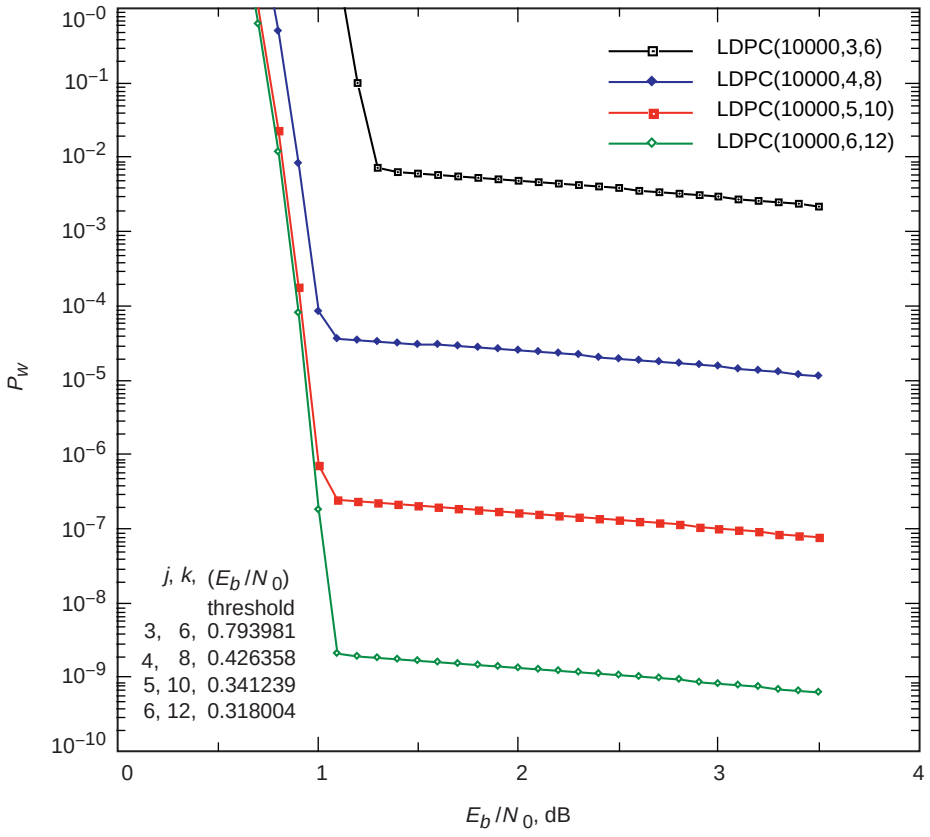

Fig. 6. Performance of low-density parity-check codes using the simple bound.

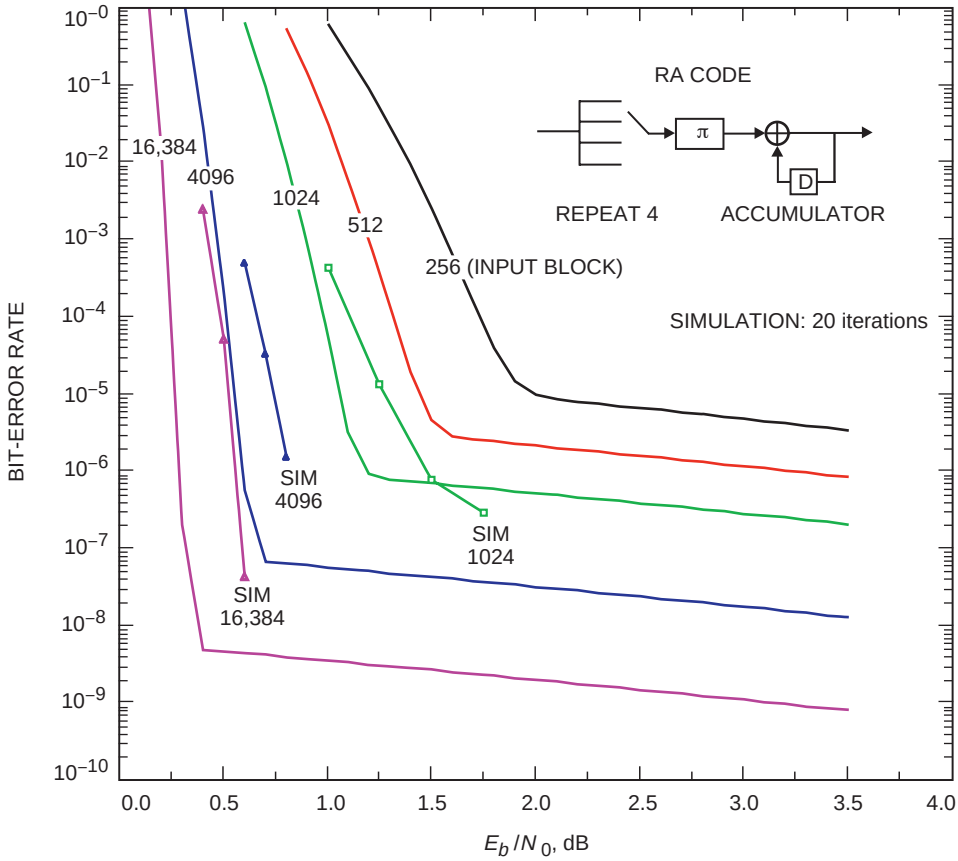

Fig. 7. ML upper bound on the bit-error probability of rate-1/4 RA codes using the simple closed-form bound, and the performance of the suboptimum iterative turbo decoder.

This threshold for the proposed bound for random codes with low rates is very close to the Shannon capacity limit. For nonrandom codes, this is the tightest closed-form threshold, better than the threshold that can be obtained from the nonmodified versions of the Gallager bound. This threshold can be used to show that, for turbo, serial, or turbo-like codes with interleaving gain, the probability of error goes to zero as the block size goes to infinity if $E_b/N_0 > (E_b/N_0)_{\text{threshold}}$. The method in [7] can be used to show this. Also see [28].

For the union bound, we have

$$E(c,h) = -r(\delta) + \delta c \tag{172}$$

where the minimum E_b/N_0 threshold based on the union bound can be computed as ($n \to \infty$):

$$\left(\frac{E_b}{N_0}\right)_{\text{threshold}} = \max_{\delta} \frac{r(\delta)}{\delta R_c} \tag{173}$$

which corresponds to the cutoff-rate threshold for random codes.

Acknowledgments

The comments of M. Klimesh, F. Pollara, S. Dolinar, H. Jin, and R. J. McEliece are acknowledged. I thank E. Biglieri for his comments on tightening the bounds, and A. M. Viterbi, A. J. Viterbi, I. Sason, and S. Shamai for providing me their papers and reports on the tangential sphere and Gallager bounds.

References

[1] R. G. Gallager, *Low density Parity Check Codes*, MIT Press, 1963.

[2] R. G. Gallager,"A Simple Derivation of the Coding Theorem and Some Applications," *IEEE Transactions on Information Theory*, pp. 3–18, January 1965.

[3] A. J. Viterbi and A. M. Viterbi, "An Improved Union Bound for Binary-Input Linear Codes on the AWGN Channel, with Applications to Turbo Decoding," *Proceedings of IEEE Information Theory Workshop*, San Diego, California, February 1998.

[4] A. M. Viterbi and A. J. Viterbi, "Improved Union Bound on Linear Codes for the Input-Binary AWGN Channel, with Applications to Turbo Codes," *Proceedings of IEEE International Symposium on Information Theory*, MIT, Boston, Massachusetts, August 16, 1998.

[5] A. M. Viterbi and A. J. Viterbi, "New Results on Serial Concatenated and Accumulated-Convolutional Turbo Code Performance," *Annales des Telecommunications*, Special Issue on Turbo Codes, vol. 54, no. 3–4, pp. 173–182, March–April 1999.

[6] A. J. Viterbi, A. M. Viterbi, J. Nicolas, and N. T. Sindhushayana, "Perspectives on Interleaved Concatenated Codes with Iterative Soft Output Decoding, *Proceedings of the International Symposium on Turbo Codes*, Brest, France, pp. 47–54, September 3–5, 1997.

[7] D. Divsalar, H. Jin, and R. J. McEliece,"Coding Theorems for 'Turbo-Like' Codes," 1998 Allerton Conference, September 23–25, 1998.

[8] 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, Massachusetts, August 16–21, 1998.

[9] I. Sason and S. Shamai, "Improved Upper Bounds on the ML Decoding Error Probability of Parallel and Serial Concatenated Turbo Codes via Their Ensemble Distance Spectrum," *IEEE Trans. on Information Theory*, to be published January 2000.

[10] I. Sason and S. Shamai (Shitz), "Bounds on the Error Probability of ML Decoding for Block and Turbo-Block Codes," *Annales des Telecommunications*, vol. 54, no. 3–4, pp. 183–200, March–April 1999.

[11] I. Sason and S. Shamai (Shitz), "Improved Upper Bounds on the ML Performance of Turbo Codes for Interleaved Rician Fading Channels, with Comparison to Iterative Decoding," to be presented at ICC 2000, New Orleans, Louisiana, June 18–22, 2000.

[12] I. Sason and S. Shamai (Shitz), "Improved Upper Bounds on the Ensemble Performance of ML Decoded Low Density Parity Check Codes," to be published in *IEEE Communications letters*.

[13] I. Sason and S. Shamai, "Gallager's 1963 Bound: Extensions and Observations," Technical Report, CC no. 258, Technion, Israel, October 1998. Updated version August 1999. Submitted for presentation at the IEEE 21th Conference of Electrical Engineers, April 11–12, 2000, Tel Aviv, Israel.

[14] T. M. Duman and M. Salehi, "New Performance Bounds for Turbo Codes," *IEEE Transactions on Communications*, vol. 46, no. 6, pp. 717–723, June 1998.

[15] 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, [date].

[16] E. R. Berlekamp, "The Technology of Error Correction Codes," *Proceedings of the IEEE*, vol. 68, no. 5, pp. 564–593, May 1980.

[17] B. Hughes, "On the Error Probability of Signals in Additive White Gaussian Noise," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 151–155, January 1991.

[18] S. Dolinar, L. Ekroot, and F. Pollara, "Improved Error Probability Bounds for Block Codes on the Gaussian Channel," *Proceedings of 1994 IEEE International Symposium on Information Theory*, Trondheim, Norway, June 24–July 1, 1994.

[19] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding: Turbo Codes," *Proc. 1993 IEEE International Conference on Communications*, Geneva, Switzerland, pp. 1064–1070, May 1993.

[20] D. Divsalar and F. Pollara, "Turbo Codes for PCS Applications," ICC '95, 'Gateway to Globalization,' 1995 IEEE International Conference on Communications, Seattle, Washington, vol. 1, pp. 54–59, 1995.

[21] D. Divsalar and F. Pollara, "Hybrid Concatenated Codes and Iterative Decoding," *The Telecommunications and Data Acquisition Progress Report 42-130, April–June 1997*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–23, August 15, 1997.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-130/130I.pdf
Also, http://www331.jpl.nasa.gov/public/JPLtcodes.html

[22] S. Benedetto and G. Montorsi, "Unveiling Turbo Codes: Some Results on Parallel Concatenated Coding Schemes," *IEEE Trans. Inf. Theory*, March 1996.

[23] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial Concatenation of Interleaved Codes: Performance Analysis, Design, and Iterative Decoding," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 909–926, May 1998.

[24] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Self-Concatenated Trellis Coded Modulation with Self-Iterative Decoding," IEEE Global Telecommunications Conference, Sydney, Australia, November 1998.

[25] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Analysis, Design, and Iterative Decoding of Double Serially Concatenated Codes with Interleavers," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 231–244, February 1998.

[26] C. E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[27] D. Divsalar and E. Biglieri,"Upper Bounds to Error Probabilities of Coded Systems Beyond the Cutoff Rate," to be presented at the IEEE International Symposium on Information Theory, 2000 (ISIT 2000), Sorrento, Italy, June 25–30, 2000.

[28] H. Jin and R. J. McEliece, "AWGN Coding Theorems for Serial Turbo Codes," Thirty-Seventh Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, September 22–24, 1999.

Appendix

Calculation of Input–Output Weight Coefficients for Turbo Codes

Consider an (n, k) block code constructed from a terminated convolutional code. Then the input–output weight enumerating function (IOWEF) is

$$A(W, H) = \sum_{w=0}^{k} \sum_{h=0}^{n} A_{w,h} W^w H^h \tag{A-1}$$

This is a two-dimensional Z-transform. Inverse Z-transform results in $A_{w,h}$. Calculation of $A_{w,h}$ is illustrated by means of an example (this is based on results by Viterbi et al. [6]). Consider a rate-1/2, four-state systematic recursive convolutional code and its state diagram, as shown in Fig. A-1.

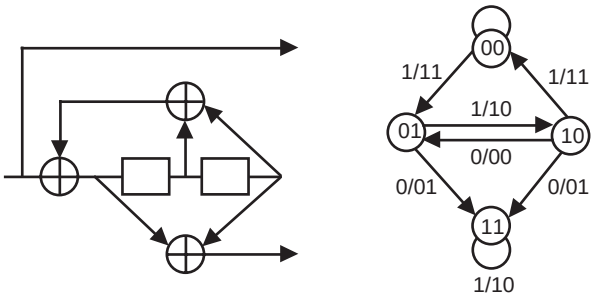

Fig. A-1. A rate-1/2, four-state systematic recursive convolutional code.

The state-transition matrix for this code is

$$C(W,H) = \begin{array}{c} \\ \text{NS} \\ \end{array} \begin{array}{c} \begin{array}{cccc} \text{PS } 00 & 01 & 10 & 11 \end{array} \\ \begin{pmatrix} 1 & 0 & WH^2 & 0 \\ WH^2 & 0 & 1 & 0 \\ 0 & WH & 0 & H \\ 0 & H & 0 & WH \end{pmatrix} \end{array} \begin{array}{l} 00 \\ 01 \\ 10 \\ 11 \end{array} \tag{A-2}$$

and the IOWEF is

$$A(W,H) = [\,1 \quad 0 \quad 0 \quad 0\,] C^k(W,H) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A-3}$$

The state-transition equations, with the initial conditions to obtain $A_{w,h}$, are

$$\begin{bmatrix} A^{(00)} & (W,H,t) \\ A^{(01)} & (W,H,t) \\ A^{(10)} & (W,H,t) \\ A^{(11)} & (W,H,t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & WH^2 & 0 \\ WH^2 & 0 & 1 & 0 \\ 0 & WH & 0 & H \\ 0 & H & 0 & WH \end{bmatrix} \begin{bmatrix} A^{(00)} & (W,H,t-1) \\ A^{(01)} & (W,H,t-1) \\ A^{(10)} & (W,H,t-1) \\ A^{(11)} & (W,H,t-1) \end{bmatrix} \tag{A-4}$$

and

$$\left.\begin{aligned} A^{(00)}_{w,h}(t) &= A^{(00)}_{w,h}(t-1) + A^{(10)}_{w-1,h-2}(t-1) \\ A^{(01)}_{w,h}(t) &= A^{(00)}_{w-1,h-2}(t-1) + A^{(10)}_{w,h}(t-1) \\ A^{(10)}_{w,h}(t) &= A^{(01)}_{w-1,h-1}(t-1) + A^{(11)}_{w,h-1}(t-1) \\ A^{(11)}_{w,h}(t) &= A^{(01)}_{w,h-1}(t-1) + A^{(11)}_{w-1,h-1}(t-1) \end{aligned}\right\} \tag{A-5}$$

with initial conditions $A^{(00)}_{0,0}(0) = 1$, $A^{(s)}_{w,h}(0) = 0$, all $(s) \neq (00)$, and $A^{(s)}_{w,h}(0) = 0$ for all (s) when $w, h < 0$. The final result is $A_{w,h} = A^{(00)}_{w,h}(k)$.

For turbo codes, suppose we want to compute

$$A_{w,h} = \sum_{h_1,h_2:h_1+h_2=h} \frac{A_{w,h_1} A_{w,h_2}}{\binom{k}{w}} \tag{A-6}$$

To prevent numerical problems we should compute

$$\tilde{A}^s_{w,h}(t) = \frac{A^s_{w,h}(t)}{\binom{t}{w}^\beta} \tag{A-7}$$

This results in a new set of difference equations:

$$\left.\begin{aligned}
\tilde{A}^{(00)}_{w,h}(t) &= f_0 \tilde{A}^{(00)}_{w,h}(t-1) + f_1 \tilde{A}^{(10)}_{w-1,h-2}(t-1) \\
\tilde{A}^{(01)}_{w,h}(t) &= f_1 \tilde{A}^{(00)}_{w-1,h-2}(t-1) + f_0 \tilde{A}^{(10)}_{w,h}(t-1) \\
\tilde{A}^{(10)}_{w,h}(t) &= f_1 \tilde{A}^{(01)}_{w-1,h-1}(t-1) + f_0 \tilde{A}^{(11)}_{w,h-1}(t-1) \\
\tilde{A}^{(11)}_{w,h}(t) &= f_0 \tilde{A}^{(01)}_{w,h-1}(t-1) + f_1 \tilde{A}^{(11)}_{w-1,h-1}(t-1)
\end{aligned}\right\} \tag{A-8}$$

where $f_0 = (1 - [w/t])^{\beta}$ and $f_1 = (w/t)^{\beta}$.