A New Upper Bound on the ML Decoding Error Probability of Linear Binary Block Codes in AWGN Interference

Shahram Yousefi and Amir K. Khandani

University of Waterloo

Coding and Signal Transmission (CST) Laboratory

Department of Electrical & Computer Engineering

Waterloo, Ontario, Canada, N2L 3G1

Technical Report UW-E&CE#2002-12

July 1st, 2002

A New Upper Bound on the ML Decoding Error Probability of Linear Binary Block Codes in AWGN Interference

Shahram Yousefi and Amir K. Khandani

Abstract

Performance evaluation of Maximum-Likelihood (ML) soft-decision-decoded binary block codes is usually carried out using bounding techniques. Many tight upper bounds on the error probability of binary codes are based on the so-called Gallager's First Bounding Technique (GFBT). The Tangential Sphere Bound (TSB) of Poltyrev which is believed for many years to offer the tightest bound developed for binary block codes is an example. Within the framework of the TSB and GFBT, we apply a new method referred to as the "Added-Hyper-Plane" (AHP) technique, to the decomposition of the error probability. This results in a bound developed upon the application of two stages of the GFBT with two different Gallager regions culminating to a tightened upper bound beyond the TSB. The proposed bound is simple and only requires the spectrum of the binary code.

Keywords

Binary linear block codes, block codes, decoding error probability, maximum-likelihood (ML) decoding, additive white Gaussian channel, bounds, upper bounds, Gallager bounds, union bound, distance spectrum.

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1 (email:shahram,khandani@shannon.uwaterloo.ca).

I. Introduction

The problem of performance evaluation of linear binary block codes with soft decision Maximum-Likelihood (ML) decoding in Additive White Gaussian Noise (AWGN) interference has long been a central problem in coding theory and practice [31]. In most of the cases, the derivation of a closed-form expression for the bit or word error probabilities is intractable. Thus, one usually resorts to bounding techniques for the aforementioned probabilities.

The most commonly used upper bound on the error probability of a digital communication system is the *union bound*. Union bound is in fact an inequality from the class of *Bonferroni-type* [12] inequalities in probability theory. These are inequalities that are universally true regardless of the underlying probability space and for all choices of the basic events. There are various Bonferroni-type upper as well as lower bounds exploited in communication theory such as KAT bound by Kuai et al. [19] (also see references in [19]). For the calculation of the union bound on the error probability of a binary block code, one only needs to have the weight enumerating function (spectrum) of the code which results in much simplicity of calculation. The union bound is quite accurate for high SNR's while for other SNR's, it is a very poor upper bound. For some applications such as concatenated coding schemes where the inner code is a binary block code, the low-SNR coding gain of the code is needed for the performance evaluation of the overall scheme which explains the need to have tighter bounds at low SNR regions. Also, for longer binary block codes ML decoding becomes prohibitively complex. Within this context, tighter upper bounds on the ML decoding performance of binary block codes used in conjunction with Binary Phase Shift Keying (BPSK) modulation will provide means of assessing the performance of these codes.

The recent overwhelming attention given to the bounding techniques for performance evaluation of codes is mainly due to the introduction of some near-Shannon-limit performing schemes. Turbo codes, invented by Berrou et al. [2] in 1993, Repeat-Accumulate (RA) codes of Divsalar et al. [6], and Low Density Parity Check (LDPC) codes of Gallager [13], resurrected by MacKay et al. [21] in 1996 are the best examples. In addition to simulations, the aforementioned schemes can be analyzed using the union bound which is a very loose measure of performance for rates above the cutoff rate of the channel [14]. This is a region which is of particular interest for long near-shannon-limit performing block codes. Therefore, there has been an increasing demand for tighter bounds on the ML decoding of such codes above the cutoff rate [8–10, 25–28].

One of the first works devoted to the performance of binary codes at low signal-to-noise ratios is that of Posner [23] which mainly revolves around quantized channel, i.e., with hard decision. A belated continuation to the work of Posner for the un-quantized channel output (soft decision) is that of Chao et. al [3]. In this work, a power series expansion of the probability of correct decision around zero SNR is used to compute a relatively accurate, albeit complex, approximation to the word error probability. The complexity of their result is due to the fact that their expression for the error probability is a function of a parameter which depends on the "global" geometrical properties of the code.

One important improvement to the union bound is that of Hughes [17]. Hughes represented the complement of the Voronoi region (all Voronoi regions are congruent to each other for Slepian codes[1] [33, 34]) as the union of a set of truncated polyhedral cones and then cleverly upper-bounded the error probability by replacing those truncated cones by truncated right circular cones with the same solid angle. For the latter the corresponding probabilities are larger but can be easily evaluated. Since the codewords of a binary code[2] used with binary modulation are not spread uniformly on the Euclidean sphere, Hughes' bound cannot be asymptotically tight. Hughes work launched a number of similar works with applications from linear binary block codes to coded modulation and concatenated codes both in AWGN and fading environments [15, 16, 22].

II. Bounds based on Gallager's first bounding technique

Many tight bounds, as noted by Divsalar [7], "essentially use a general bounding technique developed by Gallager [13]". In this method, Gallager bisects the error probability to the joint probability of error and noise residing in a region $\Re$ (here referred to as the *Gallager region*) plus joint probability of error and noise residing in the complement of $\Re$ (referred to as first and second terms, respectively); where $\Re$ is a volume around the

[1] A Slepian signal set is a Geometrically-Uniform (GU) [11] and equi-energy (sphere) signal set.

[2] In this correspondence by a binary code we mean a linear binary block code.

transmitted codeword. Divsalar [7] refers to this as "Gallager's First Bounding Technique" (GFBT). In original Gallager's work $\Re$ is a complicated region in $\mathbf{R^n}$.

For instance, the well-known Tangential Bound (TB) of Berlekamp [1] uses Gallager's first bounding technique combined with union bound to provide a significantly tightened bound than the conventional union bound at low SNR's. This is achieved by separating the radial and tangential components of the Gaussian noise with a half-space as the underlying Gallager region.

Herzberg and Poltyrev [15] use GFBT to derive a tight bound tighter than the TB of Berlekamp. $\Re$ is chosen to be a hyper-sphere and then the bound is tightened over the radius. This is referred to as the Sphere Bound (SB) of Herzberg and Poltyrev. They also apply their method to Block-Coded Modulation (BCM) schemes communicated over AWGN channel. BCM schemes involving MPSK (M-ary Phase Shift Keying) constellations are analogous to binary codes along with BPSK modulation as both are sphere signal sets, i.e., all the signal points reside on the surface of a hyper-sphere and therefore have the same energy.

The Tangential Sphere Bound (TSB) proposed for binary codes by Poltyrev [22] and for MPSK BCM schemes by Herzberg and Poltyrev [16] also uses GFBT where $\Re$ is a conical region. It is proven [16] that the TB is not tighter than the SB and the latter is also not tighter than the TSB. As a matter of fact, TSB is the *tightest* bound on the ML decoding error probability of binary block codes in AWGN interference known to-date [7, 26, 29].

The tightening of upper bounds on the ML decoding error probability of binary block codes within the format of the GFBT has been an evolutionary process: an evolution of the Gallager region from a half-space in the TB to a sphere in the SB and eventually to a cone in the TSB [36]. Yousefi and Khandani [36] observed the similarity between the Gallager regions for the TB, SB, and TSB and argued that the underlying Gallager regions for these bounds are special instances of a general n-dimensional geometrical body with azimuthal symmetry along an axis. They referred to this as a Hyper-Surface of Revolution (HSR). As such, a so-called Generalized Tangential Sphere Bound (GTSB) was proposed whose Gallager region is a generic HSR. For an equi-probable Slepian signal set, the optimum HSR is found to be a right circular cone and hence TSB is the tightest bound based upon the GTSB groundwork [36].

In the TSB method, the fact that codewords of a binary code with BPSK modulation are not uniformly distributed on a sphere is ignored and furthermore a simple union bound is applied to the calculation of the first term in the GFBT. These shortcomings provide more room for further tightening of the TSB. In this work we use the latter imperfection as the basis for our improvement. First, we propose a general upper bound applicable to any sphere signal set. The initial form of the proposed bound will require more information from the code than what the spectrum offers and hence will be complicated. We propose a method to get over this complexity and come up with a bound which is a mere function of the code spectrum at the expense of some loss of tightness. The resulting bound is still easy to evaluate and for cases under consideration it only requires optimization over one variable. This variable is the same as that of the TSB of Poltyrev.

Improvements are reported with respect to the TSB of Poltyrev for some long average binary codes (BCH codes of length 63).

III. Preliminaries

Consider a binary code $C = \{c_0, c_1, ..., c_{2^k-1}\}$ with parameters (n, k, d_{min}), to be used along with BPSK modulation (antipodal signaling) on an AWGN channel. The resulting signal set will be

$$\mathcal{S} = \{s_0, s_1, ..., s_{2^k-1}\}$$

where $s_i = \mathbf{m}(c_i) \in \mathbf{R}^n$. For $c_i = (c_{i1}, c_{i2}, ..., c_{in})$,

$$\mathbf{m}(c_i) = (m(c_{i1}), m(c_{i2}), ..., m(c_{in}))$$

where $m(\alpha) = \sqrt{E_s}(2\alpha - 1)$, $\alpha \in \{0, 1\}$, and E_s is the symbol energy[3]. The resulting signal set is a sphere and Slepian signal set. In particular for BPSK, denoting the Euclidean distance between two signal points s_i and s_j by $\delta(s_i, s_j)$ or simply δ_{ij}, we have

$$\delta_{ij}^2 = \delta^2(s_i, s_j) = \|s_i - s_j\|^2 = 4E_s d(c_i, c_j) = 4RE_b d(c_i, c_j) \tag{1}$$

where $R = k/n$ is the binary code rate, $\|.\|$ is the usual Euclidean norm, E_b is the information bit energy, and $d(,)$ is Hamming distance. Assuming AWGN interference, the

[3]Without loss of generality, E_s is chosen to be unity.

output of the channel will be a vector $r = s_i + \mathbf{n}$, where $\mathbf{n}$ is an n-dimensional vector whose elements are independent zero-mean Gaussian random variables with a variance of σ^2. Probability of word error for communicating one of 2^k messages in $\mathcal{S}$ through an AWGN channel will be:

$$P_w(E) = \sum_{i=0}^{2^k-1} P(E \mid s_i)P(s_i). \tag{2}$$

If the resulting *Geometrically-Uniform* (GU) signal set [11] is equi-probable, the ML optimum decoding rule will actually reduce to minimum Euclidean distance decoding strategy and

$$P_w(E) = P(E \mid s_i) \tag{3}$$

where s_i can be any signal point. We assume that s_0, signal corresponding to the all-zero codeword, c_0, has been transmitted.

IV. Expansion of the New bound based on the GFBT

The proposed new upper bound is primarily based on the Gallager's first bounding technique. According to the GFBT, given a transmitted signal, the word error probability can be decomposed as in

$$\begin{aligned} P_w(E) &= P\{\text{word error}, r \in \Re\} + P\{\text{word error}, r \notin \Re\} \\ &= P\{\text{E}, r \in \Re\} + P\{\text{E} \mid r \notin \Re\} \cdot P\{r \notin \Re\} \\ &\leq P\{\text{E}, r \in \Re\} + P\{r \notin \Re\} \end{aligned} \tag{4}$$

where r is the received signal vector and $\Re$, referred to as the Gallager region, is an appropriate region around the transmitted signal point. The choice of region $\Re$ is of utmost significance in this bounding method. Different choices of this region have resulted in various tight bounds in different ranges of signal-to-noise ratio. Examples of Gallager regions which have resulted in the tightest upper bounds are spheres (SB) [15] and right circular cones (TSB) [22].

The bound we are about to expand in this correspondence is developed for Slepian signal sets. For such codes, the Voronoi regions are all congruent to each other and they all include the origin of the n-space (n-dimensional space) [4, 34]. Within a general geometry, Yousefi and Khandani developed the so-called Generalized Tangential Sphere Bound

(GTSB) [36] using a generic Gallager region and proved that for sphere codes, hyper-cone indeed provides the tightest bound within the GTSB framework; thus climaxing at the TSB of Poltyrev. The new bound is similar to the GTSB and the TSB in that it is based upon multiple levels of separation of noise components from the rest of the noise vector; the first of which being the radial component of the noise.

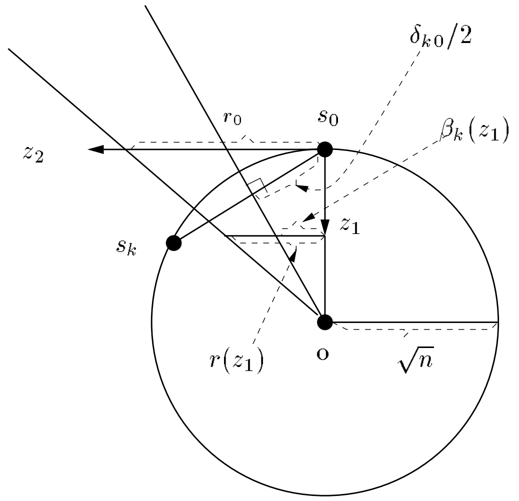


Fig. 1. Geometry of the TSB. s_0, the transmitted codeword, and s_k are both on the surface of a hyper-sphere with radius $\sqrt{n}$ centered at the origin of the space, o. For more explanation of the parameters and the geometry of this figure, see [22].

Considering the transmission of s_0 (see Fig. 1) the radial component of noise is merely the portion of the Gaussian noise along the vector $\overrightarrow{s_0 o}$.

Separating the radial component of noise, z_1, form the rest of the noise vector one can expand the word error probability $P_w(E)$ as

$$P_w(E) = \int_{-\infty}^{+\infty} P(E|z_1) f_{z_1}(z_1) dz_1 \tag{5}$$

where $f_{z_1}(z_1)$ is the zero-mean Gaussian probability density function (pdf) with a variance of σ^2.

As suggested by the TSB and GTSB, we choose the "first" Gallager region $\Re_1$ to be a hyper-cone as shown in Fig. 1. The cross sections of this hyper-cone along z_1 are hyper-spheres. As a result, for any event E, denoting $P(E|z_1)$ by $P_{z_1}(E)$, the error can be expanded as

$$P_{z_1}(E) \leq \left\{P_{z_1}(E, y \leq r^2(z_1)) + P(y > r^2(z_1))\right\} \tag{6}$$

where $y = \sum_{i=2}^{n} z_i^2$ is a random variable with Chi-square distribution with $(n-1)$ degrees of freedom [18], i.e.,

$$f_y(y) = \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})\sigma^{n-1}} \cdot e^{-\frac{y}{2\sigma^2}} y^{\frac{n-1}{2}-1} U(y) \tag{7}$$

and $\Gamma(.)$ and $U(.)$ are the complete gamma function and unit step function, respectively. With $M = 2^k - 1$, we have

$$P_{z_1}(E, y \leq r^2(z_1)) = P_{z_1}\left(\bigcup_{k=1}^{M} E_k, y \leq r^2(z_1)\right) = P_{z_1}\left(\bigcup_{k=1}^{M} \left(E_k, y \leq r^2(z_1)\right)\right) \tag{8}$$

where E_k is the error event that the received vector r is closer to s_k, than the transmitted s_0, that is,

$$E_k = \{\|r - s_k\| \leq \|r - s_0\| | s_0\} \tag{9}$$

In TSB, (8) is bounded by further separating the tangential component of noise, z_2 (where z_2 is orthogonal to z_1: $z_2 \perp z_1$), form the complete noise vector and applying a simple union bound. As a result, (6) can be seen as (see Fig. 1 and [22])

$$\begin{aligned} P_{z_1}(E) &\leq \left\{ \sum_{k:\beta_k(z_1) \leq r(z_1)} A_k P_{z_1}(E_k, y \leq r^2(z_1)) + P(y > r^2(z_1)) \right\} \\ &= \left\{ \sum_{k:d_k \leq \left\lfloor \frac{nr_0^2}{n+r_0^2} \right\rfloor} \left[A_k P(\beta_k(z_1) \leq z_2 \leq r(z_1), y_1 \leq r^2(z_1) - z_2^2)\right] + P(y > r^2(z_1)) \right\} \end{aligned} \tag{10}$$

where A_k's are the number of signal points at a Euclidean distance of $\delta_k = 2\sqrt{d_k}$ from s_0; i.e., the coefficients of the euclidean weight enumerating (ewe) function[4]

$$\mathrm{ewe}(w) = \sum_{k=1}^{n} A_k w^{\delta_k} \tag{11}$$

[4]For simplicity δ_{k0} and d_{k0} are represented by δ_k and d_k, respectively.

and $\beta_k(z_1)$, as seen from Fig. 1, is the projection of the perpendicular bisector hyper-plane between s_0 and s_k onto the $z_1 - z_2$ plane, that is, the straight line

$$\beta_k(z_1) = \frac{\sqrt{n} - z_1}{\sqrt{\frac{n}{d_k} - 1}} \tag{12}$$

and y_1 is a Chi-square distribution with $(n-2)$ degrees of freedom

$$f_{y_1}(y_1) = \frac{1}{2^{\frac{n-2}{2}}\Gamma(\frac{n-2}{2})\sigma^{n-2}} \cdot e^{-\frac{y_1}{2\sigma^2}} y_1^{\frac{n-2}{2}-1} U(y_1). \tag{13}$$

z_2 is also a zero-mean Gaussian random variable with a variance of σ^2. $\beta_k(z_1)$ is in fact the only entity in the development of the bound that solely applies to sphere constellations, hence, making the TSB limited to the equi-energy signal sets.

It is noted that in TSB we are only concerned with the upper nappe of the Gallager cone, i.e., where both $\beta_k(z_1)$ and $r(z_1)$ are nonnegative. This is a valid argument as for all practical codes and ranges of error probability the probability of being in the lower nappe is infinitesimally close to zero [26].

TSB can be further tightened by improving the tightness of the union-based upper bound in the first term of (10). Union bound is a Bonferroni-type inequality utilizing first order probabilities only. We propose the application of a Bonferroni-type inequality with degree 2 (utilizing first- and second-order probabilities).

For any set of events $E_1, E_2, ..., E_M$, we have,

$$\begin{aligned} P\left(\bigcup_{j=1}^{M} E_j\right) &= P(E_1) + P(E_2 \cap E_1^c) + ... + P(E_M \cap E_1^c \cap E_2^c \cap ... \cap E_{M-1}^c) = \\ &\sum_{j=1}^{M} P\left(E_j \bigcap \left[\bigcap_{i=1}^{j-1} E_i^c\right]\right) \end{aligned} \tag{14}$$

which can result in the second-order Bonferroni-type inequality

$$P\left(\bigcup_{j=1}^{M} E_j\right) \leq P(E_1) + P(E_2 \cap E_1^c) + P(E_3 \cap E_{j_1}^c) + ... + P(E_M \cap E_{j_{M-2}}^c) \tag{15}$$

where naturally the ordering of the events as well as the choices for the indices $j_1 \in \{1,2\}, j_2 \in \{1,2,3\}, ..., j_{M-2} \in \{1,2,...,M-1\}$ control the tightness of the bound. Denoting any of the $M!$ possible permutations of the indices of the error events E_1 to E_M by

$\Pi(1,2,\ldots,M)=(\pi_1,\pi_2,\ldots,\pi_M)$, the tightest upper bound in the form of (15) will be

$$P\left(\bigcup_{j=1}^{M} E_j\right) \leq \min_{\Pi,\Lambda}\Big\{P(E_{\pi_1})+P(E_{\pi_2}\cap E^c_{\hat{\pi}_2})+P(E_{\pi_3}\cap E^c_{\hat{\pi}_3})+\ldots \\ +P(E_{\pi_M}\cap E^c_{\hat{\pi}_M}))\Big\} \tag{16}$$

with $\Lambda=\{\hat{\pi}_2,\hat{\pi}_3,\ldots,\hat{\pi}_M\}$, in which $\hat{\pi}_j \in \{\pi_1,\pi_2,\ldots,\pi_{j-1}\}$, $j=2,3,\ldots,M$, is the index of the event that minimizes the corresponding pairwise probability. Applying (16) to (8) results in the bound

$$P_w(E) \leq \int_{-\infty}^{+\infty}\Big\{P_{z_1}\left(E_{\pi_1}, y\leq r^2(z_1)\right)+\sum_{j=2}^{M} P_{z_1}\left(E_{\pi_j}, E^c_{\hat{\pi}_j}, y\leq r^2(z_1)\right)+ \\ P(y>r^2(z_1))\Big\} f_{z_1}(z_1)dz_1 \tag{17}$$

where $\hat{\pi}_j \in \{\pi_1,\pi_2,\ldots,\pi_{j-1}\}$, $j=2,3,\ldots,M$, is the index of *error* event that minimizes the corresponding pairwise error probability.

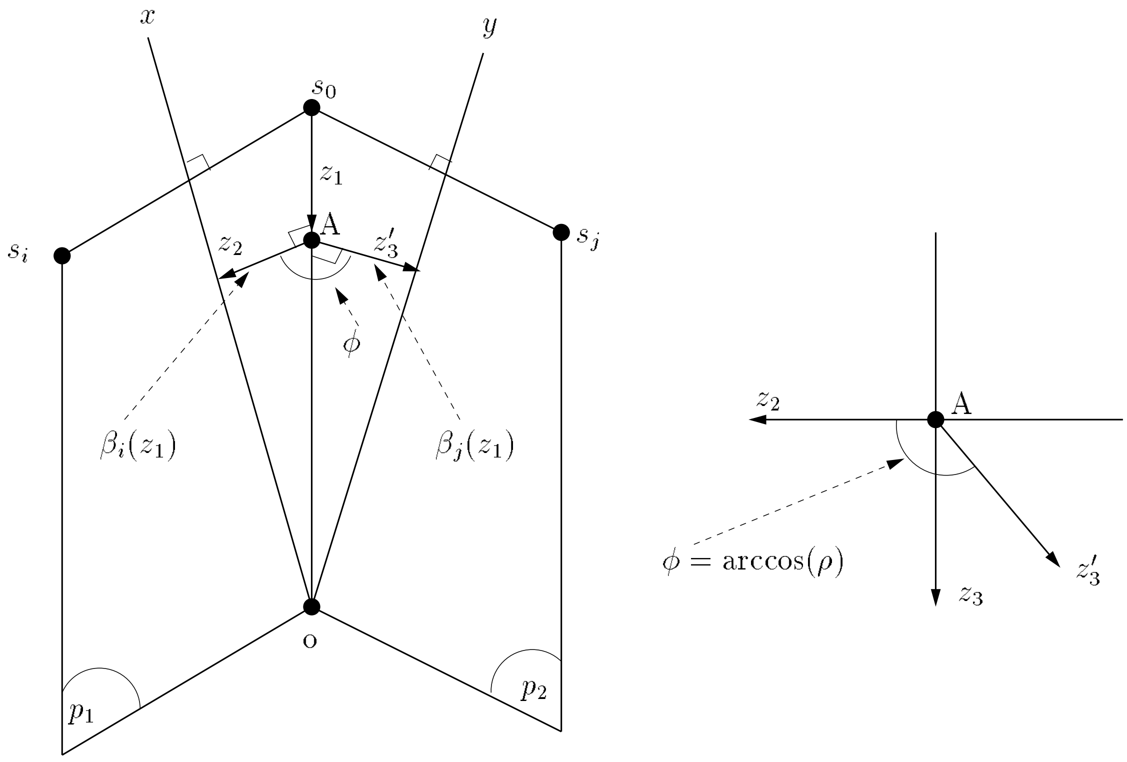
$P_{z_1}(E_{\pi_1}, y\leq r^2(z_1))$ in similarity with the TSB equals (see Fig. 1)

$$P(\beta_{\pi_1}(z_1)\leq z_2\leq r(z_1), y_1\leq r^2(z_1)-z_2^2) \tag{18}$$

where

$$\beta_{\pi_1}(z_1)=\frac{\sqrt{n}-z_1}{\sqrt{\frac{n}{d_{\pi_1}}-1}}. \tag{19}$$

For probabilities of the form $P_{z_1}(E_i, E_j^c, y\leq r^2(z_1))$ encountered in (17), we use the 3-dimensional geometry in Fig. 2-a where the transmitted point s_0 is shown in relation to two other points s_i and s_j (corresponding to error events E_i and E_j). s_0, s_i, and s_j are all on the surface of a hyper-sphere centered at the origin of the space, o, with radius $\sqrt{n}$. The 2-dimensional planes p_1 and p_2 are constructed by the points o, s_0 and s_i; and o, s_0, and s_j, respectively. These two planes hinge on z_1, the radial component of noise. The lines ox and oy are the projections of the perpendicular bisectors of lines joining s_0 to s_i and s_0 to s_j onto the planes p_1 and p_2, respectively. For sphere signal sets these perpendicular bisectors go through the origin of the signal space and therefore, so do the lines ox and oy. Then, the two noise components z_2 and z_3' on the planes p_1 and p_2 are



(a) (b)

Fig. 2. Geometry of the proposed bound in (a) 3 dimensions, and (b)in 2 dimensions.

adopted as the noise components along the directions $\overrightarrow{n_2}$ and $\overrightarrow{n_3}$ where

$$\overrightarrow{n_2} = \frac{\overrightarrow{s_0 s_i} - < \overrightarrow{s_0 s_i}, \overrightarrow{n_1} > \overrightarrow{n_1}}{\|\overrightarrow{s_0 s_i} - < \overrightarrow{s_0 s_i}, \overrightarrow{n_1} > \overrightarrow{n_1}\|} \quad , \quad \overrightarrow{n_3} = \frac{\overrightarrow{s_0 s_j} - < \overrightarrow{s_0 s_j}, \overrightarrow{n_1} > \overrightarrow{n_1}}{\|\overrightarrow{s_0 s_j} - < \overrightarrow{s_0 s_j}, \overrightarrow{n_1} > \overrightarrow{n_1}\|} \tag{20}$$

where $\overrightarrow{n_1}$ is the unit vector in the direction of z_1, and $< . >$ is the inner product operation. Then, choosing the two orthogonal noise components z_2 and z_3 appropriately (see Fig. 2 along with Appendix A) we have

$$\begin{aligned} P_{z_1}(E_i, E_j^c, y \le r^2(z_1)) = P\Big(& \beta_i(z_1) \le z_2 \le r(z_1), -r(z_1) \le z_3 \le \ell(z_1, z_2), \\ & y_2 \le r^2(z_1) - z_2^2 - z_3^2 \Big) \end{aligned} \tag{21}$$

with (see Appendix A)

$$\ell(z_1, z_2) = \frac{\beta_j(z_1) - \rho z_2}{\sin(\arccos \rho)} = \frac{\beta_j(z_1) - \rho z_2}{\sqrt{1 - \rho^2}} \tag{22}$$

and y_2 a random variable with Chi-square distribution with $(n-3)$ degrees of freedom, i.e.,

$$f_{y_2}(y_2) = \frac{1}{2^{\frac{n-3}{2}}\Gamma(\frac{n-3}{2})\sigma^{n-3}} \cdot e^{-\frac{y_2}{2\sigma^2}} y_2^{\frac{n-3}{2}-1} U(y_2). \tag{23}$$

$\rho = \cos(\phi) = \rho(z_2, z_3')$ is the correlation coefficient of z_2 and z_3' shown in Fig. 2.

Plugging (18) and (21) in (17) yields the overall bound as

$$\begin{aligned} P_w(E) \leq \min_{r(z_1),\Pi,\Lambda} \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[\int_{\beta_{\pi_1}(z_1)}^{r(z_1)} f_{z_2}(z_2) \int_0^{r^2(z_1)-z_2^2} f_{y_1}(y_1)dy_1 \cdot dz_2 + \\ \sum_{j>1:\beta_{\pi_j}(z_1)<r(z_1)} \Bigg(\int_{\beta_{\pi_j}(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell(z_1,z_2)} f_z(z_2,z_3) \int_0^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2)dy_2dz_2dz_3\Bigg) + \\ \int_{r^2(z_1)}^{+\infty} f_y(y)dy \Bigg] f_{z_1}(z_1)dz_1 \Bigg\}. \end{aligned} \tag{24}$$

The above bound is too tedious for the following reasons. Firstly, as the bound depends on the correlation coefficients between noise vectors $\overrightarrow{n_2}$ and $\overrightarrow{n_3}$, it will require the global geometrical properties of the code and only the spectrum would not suffice. Furthermore, the triple minimization over $r(z_1)$, Π, and Λ is prohibitively complex.

The choice of the optimum permutation and the optimum set Λ concurrently is tantamount to the problem of finding the optimum directed spanning graph for M nodes/vertices. This is a graph G with order $|G| = M$ and is constrained to have $\|G\| = M-1$ directed edges [24]. G is also constrained such that firstly, each vertex, except only one (referred to as the *end child*[5]), is parent to exactly one other vertex; secondly, the end child is parent to no vertex; and thirdly, edges chosen to construct the spanning graph are such that each vertex becomes a parent to either the end child or to a vertex already a parent. Each directed edge going from the parent vertex i to the child vertex j has a cost of

$$P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) \tag{25}$$

and each vertex has a cost of

$$P_{z_1}(E_i, y \leq r^2(z_1)). \tag{26}$$

[5]This is simply the first point of the constellation (with index π_1 within the formulation in (17)) selected.

The total cost of the graph constructed as such is defined as the cost of the end child plus the cost of all the edges. The graph with the above properties that has the smallest cost will be the optimum choice. Finding such a graph is a prohibitively sophisticated problem.

To detour the above-mentioned tedious obstacles and also to have a bound requiring only the spectrum of the code, we suggest the following simpler but looser bound.

For two error events E_i and E_j corresponding to codewords with Hamming weights d_i and d_j, the probability in (21) involves a correlation coefficient ρ which can not be obtained from merely Hamming weights. If one allows the second error event (E_j) to correspond to the BPSK image of a binary n-tuple of Hamming weight d_j (and not necessarily a codeword), then ρ will satisfy (see Appendix B)

$$\begin{aligned} \rho_{min} = -\min\left(\sqrt{\frac{d_i d_j}{(n-d_i)(n-d_j)}}, \sqrt{\frac{(n-d_i)(n-d_j)}{d_i d_j}}\right) \leq \rho \leq \\ \rho_{max} = \frac{\min(d_i, d_j)\left[n - \max(d_i, d_j)\right]}{\sqrt{d_i d_j (n-d_i)(n-d_j)}}. \end{aligned} \tag{27}$$

On the other hand, the probability in (21) is capped as in

$$P_{z_1}\left(E_i, E_j^c, y \leq r^2(z_1)\right) \leq P_{z_1}(E_i, E_j^c) \tag{28}$$

where the right hand side in (28) is a monotonically decreasing function of the correlation coefficient ρ in the interval $(-1, 1)$ (see Appendix C).

If A_w is the number of the codewords with a Hamming weight of w, then by adding the remaining

$$\binom{n}{w} - A_w$$

n-tuples to the codebook, an extended codebook is obtained where all the correlation coefficients (combinatorially possible) between two codewords with Hamming weights d_i and w are available. Thus, for each layer of the signal constellation, we can choose the largest ρ available with respect to an n-tuple of Hamming weight w. In this way, one only needs to find the optimum layer at which the codebook extension is done[6], i.e., finding the

[6]Instead of only at one layer/weight (w), one can also extend the codebook at multiple layers. In this work, to keep the bound simple we only use one weight for the extension.

optimum $1 \leq w \leq n$. As suggested by the TSB, we choose the curve in z_1, which specifies the boundary of the Gallager region, to be a linear function of z_1

$$r(z_1) = r_0(\sqrt{n} - z_1) \tag{29}$$

where the constant r_0 is to be optimized. With the above applied to the bound, the optimization will be over w and r_0 the choice of which are interrelated. For average codes, our iterative optimization over these two parameters indicated that the optimum r_0 is marginally the same as the radius r_0 for the TSB case [22]. The overall bound with optimum r_0 of the TSB is

$$\begin{aligned} P_w(E) \leq \min_w \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[\binom{n}{w} \int_{\beta_w(z_1)}^{r(z_1)} f_{z_2}(z_2) \int_0^{r^2(z_1)-z_2^2} f_{y_1}(y_1) dy_1 \cdot dz_2 + \\ \sum_{k:\beta_k(z_1)<r(z_1)} \Big(A_k \int_{\beta_k(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell_w^k(z_1,z_2)} f_z(z_2, z_3) \int_0^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 \Big) + \\ \int_{r^2(z_1)}^{+\infty} f_y(y) dy \Bigg] f_{z_1}(z_1) dz_1 \Bigg\} \end{aligned} \tag{30}$$

with

$$\begin{cases} \ell_w^k(z_1, z_2) = \frac{\beta_w(z_1) - \rho_{max,k} z_2}{\sqrt{1-\rho_{max,k}^2}} \\ \rho_{max,k} = \frac{\min(d_k,w)\left[n - \max(d_k,w)\right]}{\sqrt{d_k w (n-d_k)(n-w)}} \end{cases} \tag{31}$$

with,

$$1 \leq w \leq w_{max} = \left\lfloor \frac{n r_0^2}{1 + r_0^2} \right\rfloor \tag{32}$$

where w_{max} is the Hamming weight of the last layer of the code contained in the conical Gallager region.

The term in the first line of (30) corresponds to the error within the Gallager cone corresponding to the codewords (of the extended codebook) of Hamming weight w. If v_j is the BPSK image of one of a codeword of the extended codebook, then the perpendicular bisector of the line joining v_j and s_0 (which is an $(n-1)$-dimensional hyper-plane) defines an error half-space from the family

$$H_j = \{r \in R^n : \|r - v_j\| < \|r - s_0\|\} \quad , \quad j = 1, 2, \ldots, \eta = \begin{pmatrix} n \\ w \end{pmatrix} \tag{33}$$

Thus, $P(r \in H_j | s_0)$ is the probability that the received vector, r, is closer to the point v_j than the transmitted vector s_0; given the latter was transmitted. Each inequality in (33), when changed to an equality, defines a boundary hyper-plane in n-space[7].

Equivalently, the bound in (30) can be obtained by applying a second stage of the GFBT, with a "polyhedral set[8]" as its underlying Gallager region to the first term in the expansion (6) (see Appendix D).

The second Gallager region, $\Re_2$, is defined to be the following polyhedral set

$$\Re_2 = \bigcap_{j=1}^{\eta} H_j^c. \tag{34}$$

As the definition of our second Gallager region based on these hyper-planes is consequential in the derivation of the simple version of the bound in (30), we refer to the latter as the Added-Hyper-Plane (AHP) bound.

As for most codes, $A_w \ll \begin{pmatrix} n \\ w \end{pmatrix}$, the major portion of the first term in (30) is the overhead added due to the addition of the hyper-planes. The extra terms are calculated naively using simple additive (union) bound. The same methodology applied to the improvement of the TSB towards the AHP by changing probabilities of the form (26) to probabilities of the form (25) can be applied to the joint probability of error inside the cone and outside the hyper-planes corresponding to codewords (of extended codebook) at Hamming layer w. As such, the final form of the AHP bound is

[7]Note that one definition of an $(n-1)$-dimensional hyper-plane is the locus of the points in the n-space which are at the same Euclidean distance from 2 points.

[8]**Definition:** A *polyhedral set* is the intersection of a finite number of closed half-spaces [20].

$$
\begin{aligned}
P_w(E) \leq \min_{w} \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[\int_{\beta_w(z_1)}^{r(z_1)} f_{z_2}(z_2) \int_{0}^{r^2(z_1)-z_2^2} f_{y_1}(y_1) dy_1 dz_2 + \\
\left[\binom{n}{w} - 1 \right] \int_{\beta_w(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell_w(z_1,z_2)} f_z(z_2, z_3) \int_{0}^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 + \\
\sum_{k:\beta_k(z_1)<r(z_1)} \Big(A_k \int_{\beta_k(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell_w^k(z_1,z_2)} f_z(z_2, z_3) \int_{0}^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 \Big) + \\
\int_{r^2(z_1)}^{+\infty} f_y(y) dy \Bigg] f_{z_1}(z_1) dz_1 \Bigg\}
\end{aligned}
\tag{35}
$$

with (see Appendix E)

$$
\begin{cases}
\ell_w(z_1, z_2) = \frac{\beta_w(z_1) - \rho_{max,w} z_2}{\sqrt{1-\rho_{max,w}^2}} \\
\rho_{max,w} = 1 - \frac{n}{w(n-w)} \quad , \quad w \neq 0, n
\end{cases}
\tag{36}
$$

V. Examples

Our exhaustive optimization over $1 \leq w \leq \left\lfloor \frac{nr_0^2}{1+r_0^2} \right\rfloor$ showed that due to the severeness of the added overhead at low SNR's and small w's, minimum error probability is generally achieved at $w_{max} = \left\lfloor \frac{nr_0^2}{1+r_0^2} \right\rfloor$, i.e., the added hyper-planes reside in the last shell of the code within the Gallager cone. Thus, optimization is completely waived from the bound, resulting in great simplicity. In particular for longer codes such as long Turbo or LDPC codes, this provides significant reduction in the complexity of the bound.

Table I lists the optimum Gallager sphere radii for the SB (r_0^{SB} from [15], equation (12)) and the optimum r_0 for the TSB for some BCH codes of length 63.

The comparisons between the SB, TSB and AHP with $w = w_{max}$ for BCH codes of length 63 at low SNR's are presented in Figs. 3-7. For high SNR's, TSB is a very tight bound. Small but significant improvements over the TSB are observed even at very low SNR's. The significance of the small improvements achieved is more pronounced when the bound is compared to some lower bounds. For instance, for BCH$(63, 24)$ at a word error probability of 0.1 the difference between the TSB and Shannon's lower bound [30, 31] is

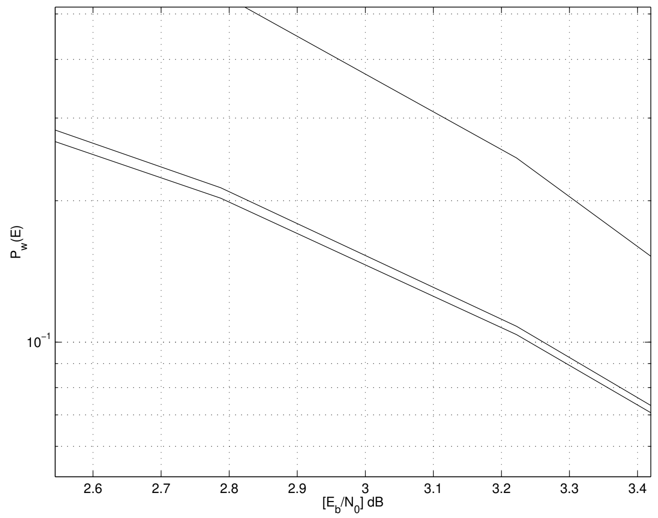


Fig. 3. Bounds on word error probability for BCH(63,51); from top to bottom: SB, TSB, and AHP with $w = 18$.

merely 0.39 dB where we report a 0.04 dB (more than 10%) tightening with the simplest form of the AHP. Shannon lower bound is shown through extensive numerical results to be the tightest lower bound on the ML decoding error probability at low SNR's [30, 35].

We have observed that a significant part of the error at high noise powers is due to the added hyper-planes in the final form of the AHP in (35). A better solution to the optimization in (24) would result in tighter bounds but will naturally be more complex. Nevertheless, any solution to these problems should merely depend on the spectrum of the code in order to keep the bound simple.

Code	$(63,24)$	$(63,30)$	$(63,36)$	$(63,39)$	$(63,51)$
r_0^{SB}	8.484	7.516	6.588	6.212	4.840
$r_0\sqrt{n}$	10.290	8.750	7.487	6.936	5.142

TABLE I

Optimum Gallager sphere radius for the SB (r_0^{SB}) and the optimum r_0 for the TSB for some BCH codes of length 63.

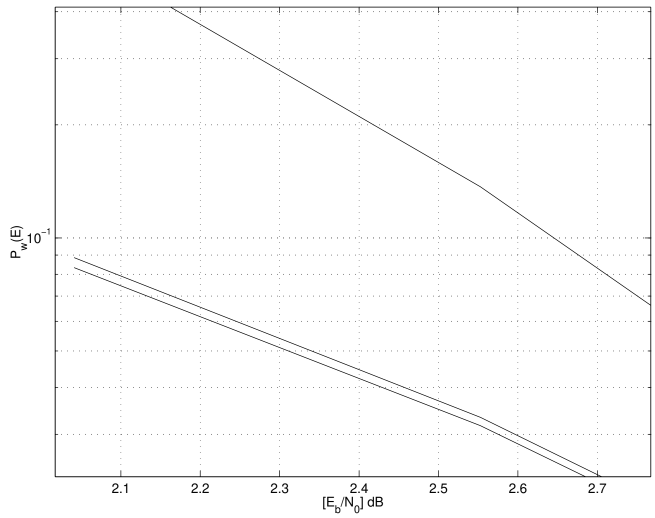


Fig. 4. Bounds on word error probability for BCH(63,39); from top to bottom: SB, TSB, and AHP with $w = 27$.

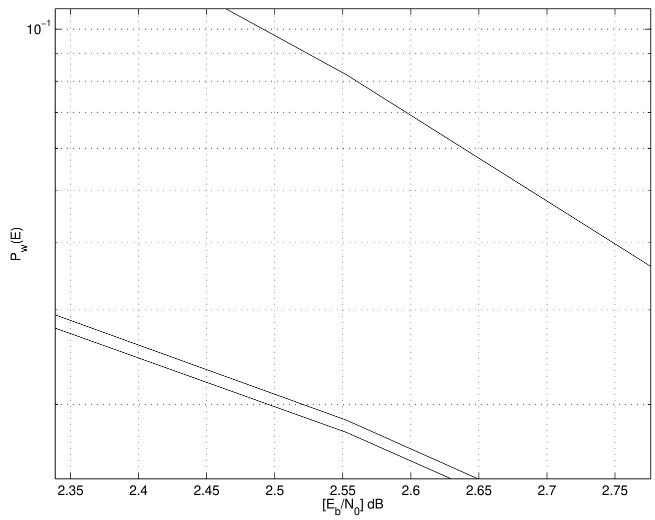


Fig. 5. Bounds on word error probability for BCH(63,36); from top to bottom: SB, TSB, and AHP with $w = 29$.

VI. Conclusion

The proposed bound in its original form in (24) is too complicated due to the triple minimization involved. Even using the same conical Gallager region as in the TSB, the optimization is still tantamount to finding an optimum directed spanning graph for a vertex set. We address the above-mentioned problems by the addition of certain hyperplanes forming a second Gallager region and thus proposing the AHP bound. AHP relies solely on the spectrum of the code and provides improvements at low SNR's with a small

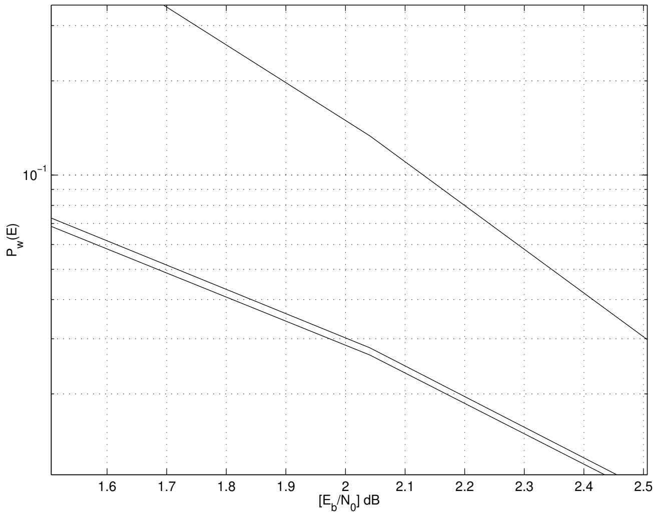


Fig. 6. Bounds on word error probability for BCH(63,30); from top to bottom: SB, TSB, and AHP with $w = 34$.

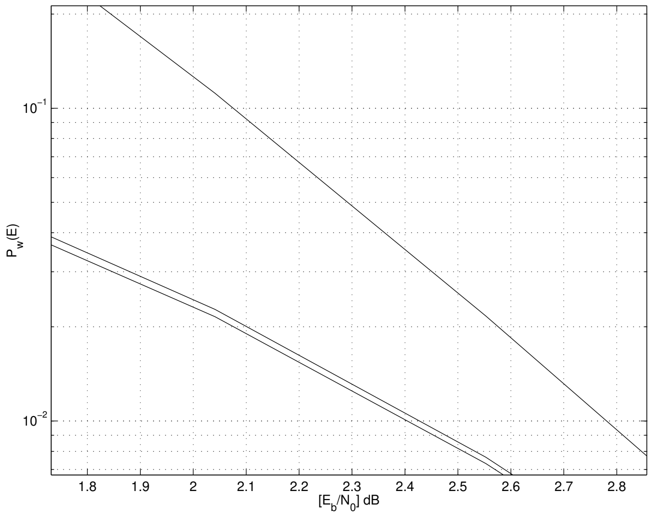


Fig. 7. Bounds on word error probability for BCH(63,24); from top to bottom: SB, TSB, and AHP with $w = 39$.

addition of complexity.

The proposed bound applies to any Slepian signal set and is not limited to the BPSK-modulated binary block codes.

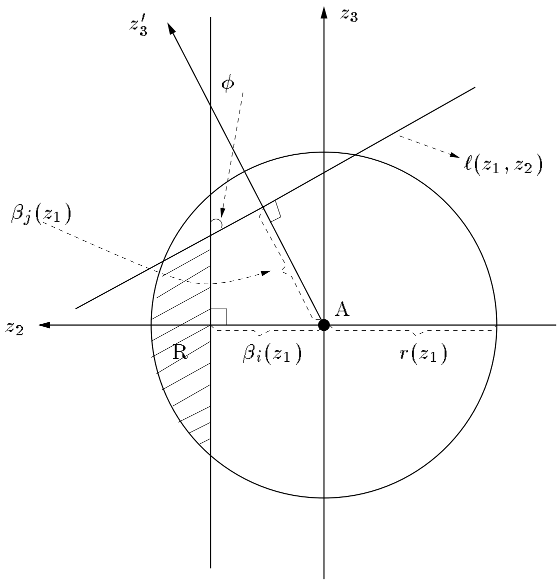


Fig. 8. Geometry for the probabilities $P_{z_1}(E_i, E_j^c, y \leq r^2(z_1))$ in the $z_2 - z_3$ plane.

Appendix A
Proof of (21) and (22)

Using the geometry of Fig. 8 in the plane of point A and noise vectors z_2 and z_3' defined in directions given in (20), we have

$$P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) = P_{z_1}\Big(\beta_i(z_1) \leq z_2, z_3' \leq \beta_j(z_1), y \leq r^2(z_1)\Big) = P(r \in \mathrm{R}) \tag{37}$$

where r is the received vector from the channel and R is the crosshatched region shown in Fig. 8. (37) in terms of two orthogonal noise components z_2 and z_3 will be

$$P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) = P\Big(\beta_i(z_1) \leq z_2 \leq r(z_1), -r(z_1) \leq z_3 \leq \ell(z_1, z_2), y_2 \leq r^2(z_1) - z_2^2 - z_3^2\Big) \tag{38}$$

where $\ell(z_1, z_2)$ is the representation for the straight line perpendicular to z_3' as shown in Fig. 8. In above and (21), z_2 and z_3 are independent zero-mean Gaussian random variables with a variance of σ^2.

Appendix B
Proof of (27)

To study the correlation coefficient

$$\rho = \rho(\overrightarrow{n_2}, \overrightarrow{n_3}) =< \overrightarrow{n_2}, \overrightarrow{n_3} >$$

we note that

$$s_0 = \mathbf{m}(c_0) = (\underbrace{-1-1\cdots-1}_{n}).$$

Notation: For a bi-valued vector $\overrightarrow{v} = (v_1, v_2, \ldots, v_n)$ with d and $(n-d)$ elements being α and γ, respectively, we use the following notation:

$$\overrightarrow{v} = \{(\alpha)^d, (\gamma)^{n-d}\} \tag{39}$$

Using the above notation, s_i and s_j are also n-tuples with ± 1 components, where the number of ones in them is d_i and d_j respectively, i.e.,

$$s_i = \{(+1)^{d_i}, (-1)^{n-d_i}\}$$

$$s_j = \{(+1)^{d_j}, (-1)^{n-d_j}\}$$

and thus,

$$\overrightarrow{s_0 s_i} = \{(2)^{d_i}, (0)^{n-d_i}\}$$

$$\overrightarrow{s_0 s_j} = \{(2)^{d_j}, (0)^{n-d_j}\}$$

Also,

$$\overrightarrow{n_1} = \frac{\overrightarrow{s_0 o}}{\|\overrightarrow{s_0 o}\|}$$

Using (20) along with the definition of $\overrightarrow{n_1}$ and straightforward manipulations we have,

$$\overrightarrow{n_2} = \frac{1}{\sqrt{\frac{d_i}{n}(n-d_i)}} \left\{(1-\frac{d_i}{n})^{d_i}, (-\frac{d_i}{n})^{n-d_i}\right\} \tag{40}$$

and

$$\overrightarrow{n_3} = \frac{1}{\sqrt{\frac{d_j}{n}(n-d_j)}} \left\{(1-\frac{d_j}{n})^{d_j}, (-\frac{d_j}{n})^{n-d_j}\right\}. \tag{41}$$

The value of $\rho =< \overrightarrow{n_2}, \overrightarrow{n_3} >$ would then depend on the location of the elements of $\overrightarrow{n_2}$ and $\overrightarrow{n_3}$ with respect to each other but it is noted that the maximum and minimum of it would correspond to maximum and minimum overlap scenarios, i.e., when the maximum/minimum number of positive elements of $\overrightarrow{n_2}$ coincide with positive elements of $\overrightarrow{n_3}$.

Then, it is straightforward to see that for maximum ρ, we have,

$$\rho_{max} = \begin{cases} \frac{d_j\left(n-d_i\right)}{\sqrt{d_i d_j (n-d_i)(n-d_j)}} = \sqrt{\frac{d_j\left(n-d_i\right)}{d_i\left(n-d_j\right)}} & \text{if} \quad d_i \geq d_j \\ \frac{d_i\left(n-d_j\right)}{\sqrt{d_i d_j (n-d_i)(n-d_j)}} = \sqrt{\frac{d_i\left(n-d_j\right)}{d_j\left(n-d_i\right)}} & \text{if} \quad d_i \leq d_j \end{cases} \tag{42}$$

and for minimum ρ, we have,

$$\rho_{min} = \begin{cases} -\sqrt{\frac{d_i d_j}{(n-d_i)(n-d_j)}} & \text{if} \quad d_i + d_j \leq n \\ -\sqrt{\frac{(n-d_i)(n-d_j)}{d_i d_j}} & \text{if} \quad d_i + d_j \geq n \end{cases}. \tag{43}$$

(42) and (43) together are equivalent to (27).

Appendix C

Proof of negative rate of change of $P(E_i, E_j^c)$ with respect to ρ

W need to prove that the function

$$f(a,b;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_a^{+\infty} \int_{-\infty}^{b} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dx dy \qquad a,b > 0 \tag{44}$$

is a monotonically decreasing function of $\rho \in (-1, +1)$.

This is done by observing that

$$f(a,b;\rho) = Q(a) - g(a,b;\rho)$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-u^2/2} du$$

is the probability under the tail of a Gaussian distribution, and

$$g(a,b;\rho) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \int_a^{+\infty} \int_b^{+\infty} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dx dy \qquad a,b > 0 \tag{45}$$

is the normalized (unit variance) two-dimensional Gaussian probability integral [32]. Function $g(a, b; \rho)$ is a monotonically increasing function of $\rho \in (-1, +1)$. To see this, $g(.)$ can be expanded to the power series [5, 30]

$$g(a, b; \rho) = Q(a)Q(b) + \frac{e^{-(a^2+b^2)}}{2\pi} \sum_{i=0}^{+\infty} \frac{H_i(-a)H_i(-b)}{(i+1)!} \rho^{i+1} \tag{46}$$

where $H_i(.)$ is the Hermite polynomial of degree i [5]. Differentiating the above with respect to ρ will yield

$$\frac{\partial}{\partial \rho} g(a, b; \rho) = \frac{e^{-(a^2+b^2)}}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{(\rho a)^2 + 2\rho a b + (\rho b)^2}{2(1-\rho^2)}\right] \tag{47}$$

which is a nonnegative quantity.

Appendix D:

AHP obtained by the application of a second stage of the GFBT

Applying the GFBT with

$$\Re_2 = \bigcap_{j=1}^{\eta} H_j^c$$

as the second Gallager region, to $P_{z_1}(E, y \leq r^2(z_1))$, we have

$$\begin{aligned} P_{z_1}(E, y \leq r^2(z_1)) &\leq P_{z_1}(E, y \leq r^2(z_1), r \in \Re_2) + P_{z_1}(y \leq r^2(z_1), r \notin \Re_2) \\ &= P_{z_1}(E, y \leq r^2(z_1), r \in \bigcap_{j=1}^{\eta} H_j^c) + P_{z_1}(y \leq r^2(z_1), r \in \bigcup_{j=1}^{\eta} H_j) \end{aligned} \tag{48}$$

We bound each of the terms in the above separately. First,

$$\begin{aligned} P_{z_1}(y \leq r^2(z_1), r \in \bigcup_{j=1}^{\eta} H_j) &= P_{z_1}(\bigcup_{j=1}^{\eta}(r \in H_j, y \leq r^2(z_1))) \\ &\leq \sum_{j=1}^{\eta} P_{z_1}(r \in H_j, y \leq r^2(z_1)) \\ &= \eta P(\beta_w(z_1) \leq z_2 \leq r(z_1), y_1 \leq r^2(z_1) - z_2^2) \end{aligned} \tag{49}$$

with

$$\beta_w(z_1) = \frac{\sqrt{n} - z_1}{\sqrt{\frac{n}{w} - 1}}. \tag{50}$$

The last equality in (49) is simply due to the SB with a geometry similar to that of Fig. 1 and expressions in (10).

Also,

$$\begin{aligned} P_{z_1}(E, y \leq r^2(z_1), r \in \bigcap_{j=1}^{\eta} H_j^c) &= P_{z_1}\Big(\bigcup_{k=1}^{M} E_k, y \leq r^2(z_1), r \in \bigcap_{j=1}^{\eta} H_j^c\Big) \\ &= P_{z_1}\Big(\bigcup_{k=1}^{M}(E_k, y \leq r^2(z_1), r \in \bigcap_{j=1}^{\eta} H_j^c)\Big) \\ &\leq P_{z_1}\Big(\bigcup_{k=1}^{M}(E_k, y \leq r^2(z_1), r \in H^c_{j^k_{opt}})\Big) \\ &\leq \sum_{k=1}^{M} P_{z_1}\Big(E_k, y \leq r^2(z_1), r \in H^c_{j^k_{opt}}\Big) \end{aligned} \tag{51}$$

where the union $\bigcap_{j=1}^{\eta} H_j^c$ is replaced by only one region, namely, $H^c_{j^k_{opt}}$, where $j^k_{opt} \in \{1, 2, \ldots, \eta\}$ can be chosen to minimize the upper bound.

Appendix E

Calculation of $\rho_{max,w}$ in (36)

In the proposed AHP upper bound, the second Gallager region involves all the codewords (in the extended codebook) of weight w. Thus, to calculate the probability $P_{z_1}(E, y \leq r^2(z_1))$, we deal with probabilities of the form $P_{z_1}(E_i, E_j^c, y \leq r^2(z_1))$ where both of the events E_i and E_j correspond to n-tuples of weight w and are chosen such that they differ from each other minimally, i.e., have a Hamming distance of 2. As such, $\rho_{max,w}$ is the maximum correlation coefficient achievable between two vectors of the form

$$\frac{1}{\sqrt{\frac{w}{n}(n-w)}}\left\{(1-\frac{w}{n})^w, (-\frac{w}{n})^{n-w}\right\}$$

It is trivial to see that the maximum correlation will be

$$\rho_{max,w} = 1 - \frac{n}{w(n-w)}.$$

References

[1] E. R. Berlekamp, "The technology of error-correcting codes," *Proc. IEEE,* vol. 68, pp. 564-593, May 1980.

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

[3] C.-C. Chao, R. J. McEliece, L. Swanson, and E. R. Rodemich, "Performance of binary block codes at low signal-to-noise ratios," *IEEE Trans. Inform. Theory*, vol. IT-38, no. 6, pp. 1677-1687, Nov. 1992.

[4] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups.* New York, NY: Springer-Verlag, 1988.

[5] H. Cramér, *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1946.

[6] D. Divsalar, H. Jin, and R. J. McEliece, " Coding theorems for 'Turbo-like' codes," *1998 Allerton Conference*, Sept. 23-25, 1998.

[7] D. Divsalar, "A simple tight bound on error probability of block codes with application to Turbo codes," *TMO Progress Report 42-139*, NASA, JPL, Pasadena, CA, USA, 1999.

[8] T. M. Duman and M. Salehi, "New performance bounds for Turbo codes," *IEEE Trans. Commun.*, vol. COM-46, no. 6, pp. 717-723, June 1998.

[9] T. M. Duman and M. Salehi, "Performance bounds for Turbo-coded modulation systems," *IEEE Trans. Commun.*, vol. COM-47, no. 4, pp. 511-521, Apr. 1999.

[10] T. M. Duman and M. Salehi, "The union bound for Turbo-coded modulation systems over fading channels," *IEEE Trans. Commun.*, vol. COM-47, no. 10, pp. 1495-1502, Oct. 1999.

[11] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. IT-37, no. 5, pp. 1241-1260, Sept. 1991.

[12] J. Galambos and I. Simonelli, *Bonferroni-type Inequalities with Applications.* New York, NY: Springer, 1996.

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

[14] R. G. Gallager, *Information Theory and Reliable Communication.* New York, NY: John Wiley & Sons, 1968.

[15] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block-coded modulation structures," *IEEE Trans. Inform. Theory*, vol. IT-40, no. 3, pp. 903-911, May. 1994.

[16] H. Herzberg and G. Poltyrev, "The error probability of M-ary PSK block-coded modulation schemes," *IEEE Trans. Commun.*, vol. COM-44, no. 4, pp. 427-433, Apr. 1996.

[17] B. Hughes, "On the error probability of signals in additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-37, no. 1, pp. 151-155, Jan. 1991.

[18] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions.* vol. 1, New York, NY: John Wiley & Sons, 1994.

[19] H. Kuai, F. Alajaji, and G. Takahara, "Tight error bounds for nonuniform signaling over AWGN Channels," *IEEE Trans. Inform. Theory*, vol. IT-46, no. 7, pp. 2712-2718, Nov. 2000.

[20] P. McMullen and G. C. Shephard, *Convex Polytopes and the Upper Bound Conjecture.* Cambridge University Press, 1971.

[21] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Lett.,* vol. 32, pp. 1645-1646, Aug. 1996.

[22] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inform. Theory*, vol. IT-40, no. 4, pp. 1284-1292, July. 1994.

[23] E. C. Posner, "Properties of error-correcting codes at low signal-to-noise ratios," *SIAM J. Appl. Math.,* vol. 15, pp. 775-798, July 1967.

[24] R. Diestel, *Graph Theory.* New York, NY: Springer-Verlag, 1997.

[25] I. Sason, "Upper bounds on maximum-likelihood decoding error probability for block codes and turbo-like codes," PhD dissertation, Department of Electrical Engineering, Technion, Israel Institute of Technology, Haifa, Israel, Sept. 2001. Available in: *http://tiger.technion.ac.il/∼eeigal/*.

[26] I. Sason and S. Shamai(Shitz), "Improved upper bounds on the decoding error probability of parallel and serial concatenated Turbo codes via their ensemble distance spectrum," *IEEE Trans. Inform. Theory*, vol. IT-46, no. 1, pp. 1-23, Jan. 2000.

[27] —, "Improved upper bounds on the ensemble performance of ML decoded low density parallel check codes," *IEEE Commun. Lett.,* vol. 4, pp. 89-91, Mar. 2000.

[28] —, "On improved bounds on the decoding error probability of block codes over interleaved fading channels, with applications to Turbo-like codes," *IEEE Trans. Inform. Theory*, vol. IT-47, no. 6, pp. 2275-2299, Nov. 2001.

[29] —, "Variations on the Gallager bounds, connections and applications," to be published in *IEEE Trans. Inform. Theory*, Available in: *http://lthiwww.epfl.ch/∼eeigal/coding-theory.html/*.

[30] G. E. Séguin, "A lower bound on the error probability for signals in white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-44, no. 7, pp. 3168-3175, Nov. 1998.

[31] C. E. Shannon, "Probability of error for optimal codes in Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611-656; and *Math Rev.*, vol. 21, p. 1920, 1959.

[32] M. K. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Trans. Commun.*, vol. COM-46, no. 2, pp. 200-210, Feb. 1998.

[33] D. Slepian, "A class of binary signaling alphabets," *Bell Sys. Tech. Journ.,* vol. 35, pp. 203-234, Jan. 1956.

[34] D. Slepian, "Group codes for the Gaussian channel," *Bell Syst. Tech. J.,* vol. 47, no. 4, pp. 572-602, Apr. 1968.

[35] P. F. Swaszek, "A lower bound on the error probability for signals in white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-41, no. 3, pp. 837-841, May 1995.

[36] S. Yousefi and A. K. Khandani, "Generalized Tangential Sphere Bound on the ML Decoding Error Probability of Linear Binary Block Codes in AWGN Interference," preprint, submitted to *IEEE Trans. Inform. Theory*, Feb. 2002.