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

Shahram Yousefi, *Member, IEEE,* and Amir K. Khandani, *Member, IEEE*

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.

Index Terms

Additive White Gaussian Noise (AWGN) channel, block codes, bounds, decoding error probability, distance spectrum, Gallager bounds, linear binary block codes, maximum-likelihood (ML) decoding, probability of error, union bound, upper bounds.

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 [1]. 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* [2] 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. [3] (also see references in [3]–[6]). For the calculation of the union bound on the error probability of a binary code[1], 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. Besides, as Shannon noted

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and was partly presented at the International Symposium on Telecommunications (IST 2003), Isfahan, Iran, Aug. 2003. S. Yousefi was with the Department of Electrical and Computer Engineering, University of Waterloo. He is now with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada K7L 3N6 (email:yousefi@ee.queensu.ca). A. K. Khandani is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1 (email:khandani@shannon.uwaterloo.ca).

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

[1], [4] the performance study of block codes is worthwhile regardless of the immediate application in mind; having made it a classical problem for decades.

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. [7] in 1993, Repeat-Accumulate (RA) codes of Divsalar et al. [8], and Low Density Parity Check (LDPC) codes of Gallager [9], resurrected by MacKay et al. [10] 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 [11]. 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 [5], [12]–[17].

One of the first works devoted to the performance of binary codes at low signal-to-noise ratios is that of Posner [18] 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 [19]. 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 [20]. Hughes represented the complement of the Voronoi region (all Voronoi regions are congruent to each other for Slepian codes[2] [21], [22]) 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 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 [24]–[26].

II. Bounds based on Gallager’s first bounding technique

Many tight upper bounds, as noted by Divsalar [27], “essentially use a general bounding technique developed by Gallager [9]”. 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 $\Re$, $\Re^c$ (referred to as first and second terms, respectively); where $\Re$ is a volume around the transmitted codeword. For obvious reasons, the regions $\Re$ and $\Re^c$ are sometimes referred to as the regions of *many* and *few* errors. Divsalar [27] 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 [28] 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 [24] 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

[2] **Definition**: A Slepian signal set in $\mathbf{R}^n$ is the orbit of a point in $\mathbf{R}^n$ under a finite group of orthogonal transformations of $\mathbf{R}^n$. As such, Slepian signal sets exhibit strong symmetry properties. All points of a Slepian signal set are exactly equivalent in every aspect except for their absolute position in the Euclidean space. From the above, it is obvious that the points of a Slepian signal set are all on the surface of a sphere (hyper-sphere) in the Euclidean n-space centered at the origin. A Slepian signal set is therefore a Geometrically-Uniform (GU) [23] and equi-energy (sphere) signal set.

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 [26] and for MPSK BCM schemes by Herzberg and Poltyrev [25] also uses GFBT where $\Re$ is a conical region. It is proven [25] analytically that union bound is not tighter than the TB and the latter is also not tighter than the TSB. Also, through extensive numerical analysis, it has been shown and is well established today that the SB is 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 [5], [6], [15], [27], [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 [30]. Yousefi and Khandani [30] 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 [30].

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, viz., within the region of many errors. 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. It is to be noted that as the conical regions employed in the development of the TSB do not form a dense packing of the Euclidean space one cannot expect the TSB to be asymptotically tight.

The rest of this article is organized as follows. We start with preliminaries and notations in section III. Following, 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. In section IV, we provide a few examples. Improvements are reported with respect to the TSB of Poltyrev for some average binary codes (BCH codes of length 63). These codes are long enough to render the ML decoding prohibitively complex while they have a spectrum which is very close to the average spectrum. The paper in concluded in section V with some discussions and remarks on the newly developed bounds in comparison with some other known bounds.

III. Preliminaries

Consider a binary code $C = \{\mathbf{c}_0, \mathbf{c}_1, ..., \mathbf{c}_{2^k-1}\}$ 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

$$\mathcal{S} = \{\mathbf{s}_0, \mathbf{s}_1, ..., \mathbf{s}_{2^k-1}\}$$

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

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

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 Slepian signal set. In particular for BPSK, denoting the Euclidean distance between two signal points $\mathbf{s}_i$ and $\mathbf{s}_j$ by $\delta(\mathbf{s}_i, \mathbf{s}_j)$ or simply δ_{ij}, we have

$$\delta_{ij}^2 = \delta^2(\mathbf{s}_i, \mathbf{s}_j) = \|\mathbf{s}_i - \mathbf{s}_j\|^2 = 4E_s d(\mathbf{c}_i, \mathbf{c}_j) = 4RE_b d(\mathbf{c}_i, \mathbf{c}_j) \tag{1}$$

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

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 output of the channel will be a vector $\mathbf{r} = \mathbf{s}_i + \mathbf{n}$, 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:

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

If the resulting *Geometrically-Uniform* (GU) signal set [23] is equi-probable, the optimum ML decoding rule will actually reduce to minimum Euclidean distance decoding strategy and

$$P_w(E) = P(E \mid \mathbf{s}_i) \tag{3}$$

where $\mathbf{s}_i$ can be any signal point. We assume that $\mathbf{s}_0$, signal corresponding to the all-zero codeword, $\mathbf{c}_0$, 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

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

where $\mathbf{r}$ 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) [24] and right circular cones (TSB) [26].

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) [22], [31]. Within a general geometry, Yousefi and Khandani developed the so-called Generalized Tangential Sphere Bound (GTSB) [30] 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.

Considering the transmission of $\mathbf{s}_0$ (see Fig. 1) the radial component of noise is merely the portion of the Gaussian noise along the vector $\overrightarrow{\mathbf{s}_0\mathbf{o}}$.

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

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

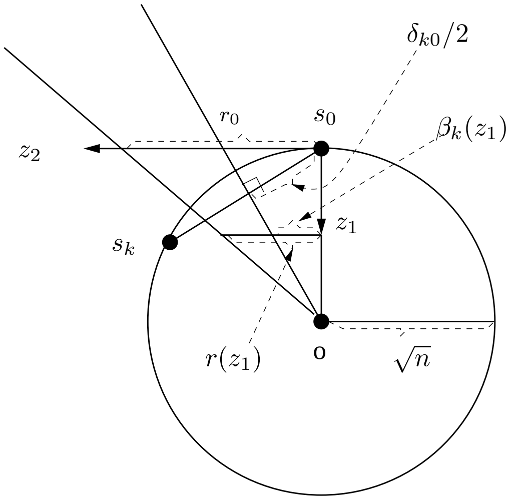

Fig. 1. Geometry of the TSB. s_0, the transmitted codeword, and $\mathbf{s}_k$ are both on the surface of a hyper-sphere with radius $\sqrt{n}$ centered at the origin of the space, **o**. For more explanation of the parameters and the geometry of this figure, see [26].

where $y = \sum_{i=2}^{n} z_i^2$ is a random variable with Chi-square distribution with $(n-1)$ degrees of freedom [32], i.e.,

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

in which $\Gamma(.)$ and $U(.)$ are the complete gamma and the unit step functions, 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 the pairwise error event E_k is the error event that the received vector $\mathbf{r}$ is closer to s_k, than the transmitted s_0, that is,

$$E_k = \{\|\mathbf{r} - \mathbf{s}_k\| \leq \|\mathbf{r} - \mathbf{s}_0\| | \mathbf{s}_0\}. \tag{9}$$

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 [26])

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

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]

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

[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 of the line joining $\mathbf{s}_0$ and $\mathbf{s}_k$ 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 [15].

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,

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

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

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

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. The above independently-developed upper bound of ours has been previously reported by Hunter in [33] in the following equivalent form

$$P\left(\bigcup_{j=1}^{M} E_j\right) \leq \min_{\Pi,\Lambda} \left\{ \sum_{j=1}^{M} P(E_{\pi_j}) - \sum_{j=2}^{M} P(E_{\pi_j} \cap E_{\hat{\pi}_j}) \right\}. \tag{17}$$

Applying (16) to (8) results in the bound

$$P_w(E) \leq \int_{-\infty}^{+\infty} \Big\{ P_{z_1}\big(E_{\pi_1}, y \leq r^2(z_1)\big) + \sum_{j=2}^{M} P_{z_1}\big(E_{\pi_j}, E^c_{\hat{\pi}_j}, y \leq r^2(z_1)\big) + P(y > r^2(z_1)) \Big\} f_{z_1}(z_1) dz_1 \tag{18}$$

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)

$$P(\beta_{\pi_1}(z_1) \leq z_2 \leq r(z_1), y_1 \leq r^2(z_1) - z_2^2) \tag{19}$$

where

$$\beta_{\pi_1}(z_1) = \frac{\sqrt{n} - z_1}{\sqrt{\frac{n}{d_{\pi_1}} - 1}}. \tag{20}$$

For probabilities of the form $P_{z_1}(E_i, E_j^c, y \leq r^2(z_1))$ encountered in (18), 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 $\mathrm{o}x$ and $\mathrm{o}y$ are the projections of the perpendicular bisectors of lines joining s_0 to s_i

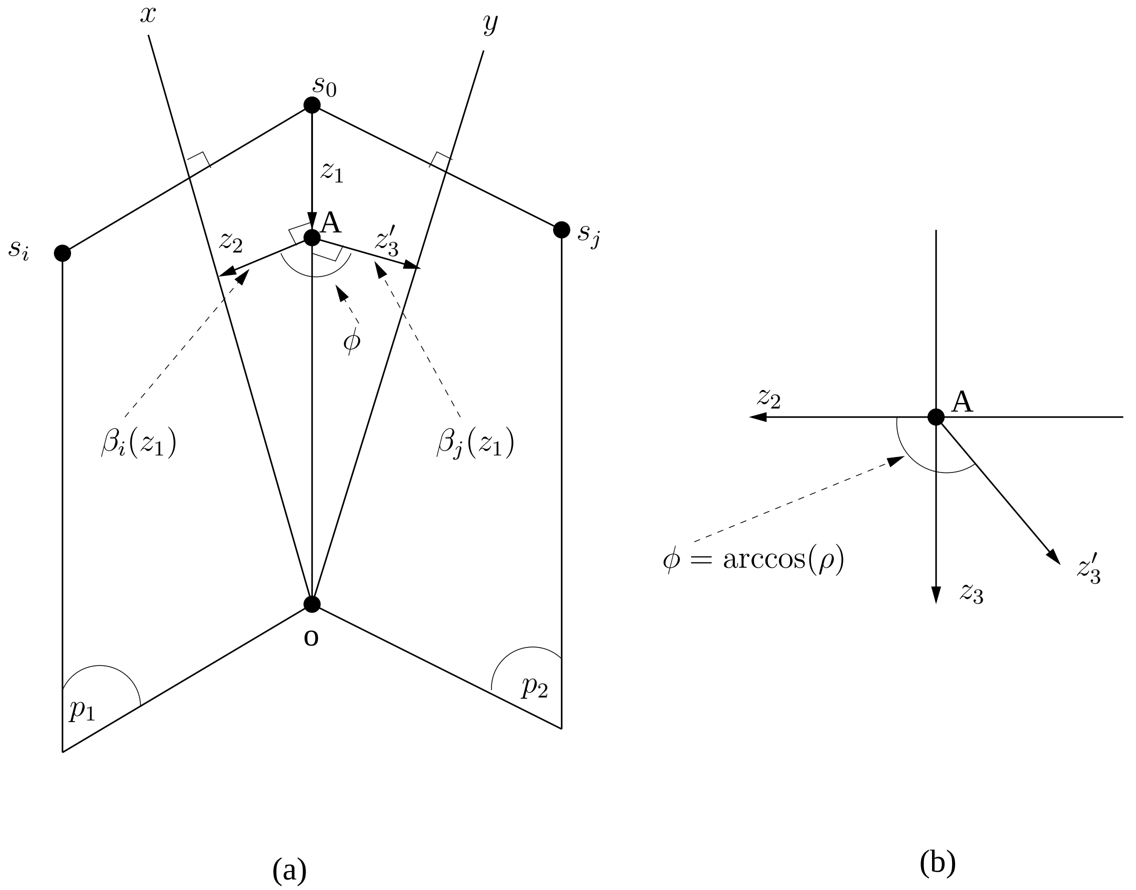

Fig. 2. Geometry of the proposed bound in (a) 3 dimensions, and (b)in 2 dimensions.

and $\mathbf{s}_0$ to $\mathbf{s}_j$ 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 adopted as the noise components along the directions $\overrightarrow{n_2}$ and $\overrightarrow{n_3}$ where

$$\overrightarrow{n_2} = \frac{\overrightarrow{s_0 s_i} - < \overrightarrow{s_0 s_i}, \overrightarrow{n_1} > \overrightarrow{n_1}}{\|\overrightarrow{s_0 s_i} - < \overrightarrow{s_0 s_i}, \overrightarrow{n_1} > \overrightarrow{n_1}\|} \quad , \quad \overrightarrow{n_3} = \frac{\overrightarrow{s_0 s_j} - < \overrightarrow{s_0 s_j}, \overrightarrow{n_1} > \overrightarrow{n_1}}{\|\overrightarrow{s_0 s_j} - < \overrightarrow{s_0 s_j}, \overrightarrow{n_1} > \overrightarrow{n_1}\|} \tag{21}$$

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

$$\begin{aligned} P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) = P\Big(&\beta_i(z_1) \leq z_2 \leq r(z_1), -r(z_1) \leq z_3 \leq \ell(z_1, z_2), \\ &y_2 \leq r^2(z_1) - z_2^2 - z_3^2\Big) \end{aligned} \tag{22}$$

with (see Appendix A)

$$\ell(z_1, z_2) = \frac{\beta_j(z_1) - \rho z_2}{\sin(\arccos \rho)} = \frac{\beta_j(z_1) - \rho z_2}{\sqrt{1 - \rho^2}} \tag{23}$$

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

$$f_{y_2}(y_2) = \frac{1}{2^{\frac{n-3}{2}} \Gamma(\frac{n-3}{2}) \sigma^{n-3}} \cdot e^{-\frac{y_2}{2\sigma^2}} y_2^{\frac{n-3}{2}-1} U(y_2). \tag{24}$$

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

Plugging (19) and (22) in (18) yields the overall bound as

$$\begin{aligned} P_w(E) \leq \min_{r(z_1),\Pi,\Lambda} \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[&\int_{\beta_{\pi_1}(z_1)}^{r(z_1)} f_{z_2}(z_2) \int_0^{r^2(z_1)-z_2^2} f_{y_1}(y_1) dy_1 \cdot dz_2 + \\ \sum_{j>1:\beta_{\pi_j}(z_1)<r(z_1)} \Bigg(&\int_{\beta_{\pi_j}(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell(z_1,z_2)} f_z(z_2, z_3) \int_0^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 \Bigg) + \\ &\int_{r^2(z_1)}^{+\infty} f_y(y) dy \Bigg] f_{z_1}(z_1) dz_1 \Bigg\}. \end{aligned} \tag{25}$$

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 [34]. 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

[5] This is simply the first point of the constellation (with index π_1 within the formulation in (18)) selected.

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

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

and each vertex has a cost of

$$P_{z_1}(E_i, y \leq r^2(z_1)). \tag{27}$$

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. Hunter [33] proves systematically that any spanning tree[6] of the M vertices of events can be used in (17), and thus also in our (16) and (25). In [33], Kruskal's [35] minimal spanning tree algorithm has been proposed to tighten the upper bound. The other possibility is the well known greedy minimal spanning tree algorithm of Prim [36]. In both cases the local and global optima coincide [37], [38]. In smaller problems such as the uncoded modulation case considered in [3] such greedy algorithms are very efficient but in cases like the problem in this article, the small improvements obtained are well canceled out by the significant algorithm overhead. This is even more pronounced here as the bounds in this paper are intended for longer and larger codes and high-dimensional constellations.

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 (22) 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)

$$\begin{aligned} \rho_{min} = -\min\left(\sqrt{\frac{d_i d_j}{(n-d_i)(n-d_j)}}, \sqrt{\frac{(n-d_i)(n-d_j)}{d_i d_j}}\right) \leq \rho \leq \\ \rho_{max} = \frac{\min(d_i,d_j)\big[n - \max(d_i,d_j)\big]}{\sqrt{d_i d_j (n-d_i)(n-d_j)}}. \end{aligned} \tag{28}$$

On the other hand, the probability in (22) is capped as in

$$P_{z_1}\Bigg(E_i, E_j^c, y \leq r^2(z_1)\Bigg) \leq P_{z_1}(E_i, E_j^c) \tag{29}$$

where the right hand side in (29) is a monotonically decreasing function of the correlation coefficient ρ in the interval $(-1, 1)$. This is obvious as it has been proven that probabilities of the form $P_{z_1}(E_i, E_j)$ are monotonically increasing with the correlation coefficient [39].

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

$$\left(\begin{array}{c} n \\ w \end{array}\right) - A_w$$

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

[6] **Definition:** A spanning tree of M nodes is a connected graph with $(M-1)$ branches such that at least one branch is incident on any of the nodes. This means that there is exactly one path between any two vertices of the spanning tree.

extension is done[7], i.e., finding the 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

$$r(z_1) = r_0(\sqrt{n} - z_1) \tag{30}$$

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 [26]. The overall bound with optimum r_0 of the TSB is

$$\begin{aligned} P_w(E) \leq \min_w \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[\binom{n}{w} \int_{\beta_w(z_1)}^{r(z_1)} f_{z_2}(z_2) \int_0^{r^2(z_1)-z_2^2} f_{y_1}(y_1) dy_1 \cdot dz_2 + \\ \sum_{k:\beta_k(z_1)<r(z_1)} \Bigg(A_k \int_{\beta_k(z_1)}^{r(z_1)} \int_{-r(z_1)}^{\ell_w^k(z_1,z_2)} f_z(z_2, z_3) \int_0^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 \Bigg) + \\ \int_{r^2(z_1)}^{+\infty} f_y(y) dy \Bigg] f_{z_1}(z_1) dz_1 \Bigg\} \end{aligned} \tag{31}$$

with

$$\begin{cases} \ell_w^k(z_1, z_2) = \frac{\beta_w(z_1) - \rho_{max,k} z_2}{\sqrt{1-\rho_{max,k}^2}} \\ \rho_{max,k} = \frac{\min(d_k,w)\left[n - \max(d_k,w)\right]}{\sqrt{d_k w (n-d_k)(n-w)}} \end{cases} \tag{32}$$

and,

$$1 \leq w \leq w_{max} = \left\lfloor \frac{n r_0^2}{1 + r_0^2} \right\rfloor \tag{33}$$

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 (31) corresponds to the error within the Gallager cone corresponding to the codewords (of the extended codebook) of Hamming weight w. If $\mathbf{v}_j$ is the BPSK image of one of a codeword of the extended codebook, then the perpendicular bisector of the line joining $\mathbf{v}_j$ and $\mathbf{s}_0$ (which is an $(n-1)$-dimensional hyper-plane) defines an error half-space from the family

$$H_j = \{\mathbf{r} \in R^n : \|\mathbf{r} - \mathbf{v}_j\| < \|\mathbf{r} - \mathbf{s}_0\|\} \quad , \quad j = 1, 2, \ldots, \eta = \binom{n}{w} \tag{34}$$

Thus, $P(\mathbf{r} \in H_j | \mathbf{s}_0)$ is the probability that the received vector, $\mathbf{r}$, is closer to the point $\mathbf{v}_j$ than the transmitted vector $\mathbf{s}_0$; given the latter was transmitted. Each inequality in (34), when changed to an equality, defines a boundary hyper-plane in n-space[8].

Equivalently, the bound in (31) can be obtained by applying a second stage of the GFBT, with a "polyhedral set[9]" as its underlying Gallager region to the first term in the expansion (6) (see Appendix C).

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

[7] 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.

[8] 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.

[9] **Definition:** A *polyhedral set* is the intersection of a finite number of closed half-spaces [40].

$$\Re_2 = \bigcap_{j=1}^{\eta} H_j^c. \tag{35}$$

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 (31), we refer to the latter as the Added-Hyper-Plane (AHP) bound.

As for most codes, $A_w \ll \left(\begin{array}{c} n \\ w \end{array} \right)$, the major portion of the first term in (31) 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 (27) to probabilities of the form (26) 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

$$\begin{aligned} P_w(E) \leq \min_w \Bigg\{ \int\limits_{-\infty}^{+\infty} \Bigg[\int\limits_{\beta_w(z_1)}^{r(z_1)} f_{z_2}(z_2) \int\limits_{0}^{r^2(z_1)-z_2^2} f_{y_1}(y_1) dy_1 dz_2 + \\ \left[\left(\begin{array}{c} n \\ w \end{array} \right) - 1 \right] \int\limits_{\beta_w(z_1)}^{r(z_1)} \int\limits_{-r(z_1)}^{\ell_w(z_1,z_2)} f_z(z_2,z_3) \int\limits_{0}^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 + \\ \sum_{k:\beta_k(z_1)<r(z_1)} \Bigg(A_k \int\limits_{\beta_k(z_1)}^{r(z_1)} \int\limits_{-r(z_1)}^{\ell_w^k(z_1,z_2)} f_z(z_2,z_3) \int\limits_{0}^{r^2(z_1)-z_2^2-z_3^2} f_{y_2}(y_2) dy_2 dz_2 dz_3 \Bigg) + \\ \int\limits_{r^2(z_1)}^{+\infty} f_y(y) dy \Bigg] f_{z_1}(z_1) dz_1 \Bigg\} \end{aligned} \tag{36}$$

with (see Appendix D)

$$\left\{ \begin{array}{ll} \ell_w(z_1,z_2) = \frac{\beta_w(z_1)-\rho_{max,w} z_2}{\sqrt{1-\rho_{max,w}^2}} & \\ \rho_{max,w} = 1 - \frac{n}{w(n-w)} \quad , & w \neq 0, n \end{array} \right. . \tag{37}$$

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 r_0 for the TSB, w_{max} for the AHP bound, and θ_{SLB} of Shannon Lower Bound (SLB) for some BCH codes of length 63.

The well known Shannon lower bound, which is based on matching the shape of the Voronoi region of the constellation, is more exact for low SNR values and in fact coincides with the TSB as very low SNR. SLB on the performance of ML decoding of Slepian codes is based on upper bounding the Voronoi region of the transmitted point by a spherical n-cone with a solid angle equal to the solid angle of the n-sphere divided by the number points in the constellation and for a binary code (n, k) is given by [1], [39]:

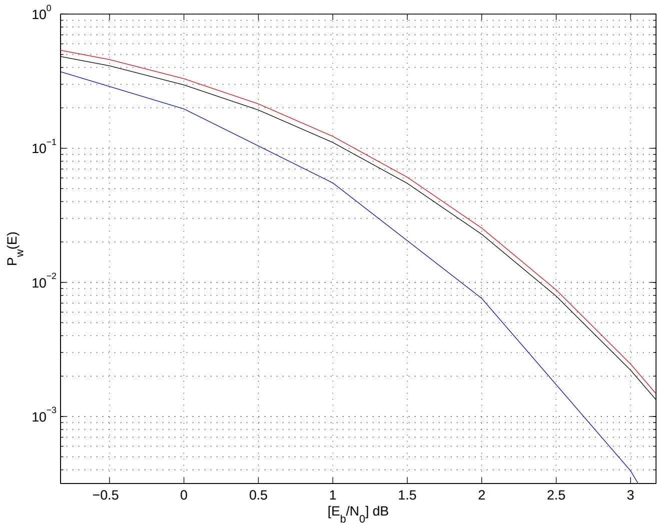

Fig. 3. Bounds on word error probability for BCH(63,24); from top to bottom: TSB, AHP with $w = 39$, and SLB.

$$P_w(E) \geq \frac{1}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)} \int\limits_0^\infty u^{n-2} e^{-\frac{u^2}{2}} Q\left(\sqrt{\frac{2nRE_b}{N_0}} - u \ \cot(\theta_{SLB})\right) du \tag{38}$$

where the optimal angle θ_{SLB} is obtained from

$$\int\limits_0^{\theta_{SLB}} (\sin u)^{n-2} du = \frac{n\sqrt{\pi}}{(n-1)2^k} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}. \tag{39}$$

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 [39], [41].

The comparisons between the TSB, AHP with $w = w_{max}$, and the SLB for two BCH codes of length 63 at low SNR's are presented in Figs. 3-4. 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. Note, for instance, that for BCH$(63, 24)$ at a word error probability of 0.1 the difference between the TSB and SLB is merely 0.59 dB where we report a 0.08 dB (more than 10%) tightening with the simplest form of the AHP.

We have observed that a significant part of the error at high noise powers is due to the added hyperplanes in the final form of the AHP in (36). A better solution to the optimization in (25) 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	1.2964	1.1024	0.9433	0.8738	0.6478
w_{max}	39	34	29	27	18
θ_{SLB}	0.9210	0.8415	0.7727	0.7414	0.6328

TABLE I
THE OPTIMUM r_0 FOR THE TSB, w FOR AHP, AND θ_{SLB} FOR SOME BCH CODES OF LENGTH 63.

VI. CONCLUSION

The proposed bound in its original form in (25) 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

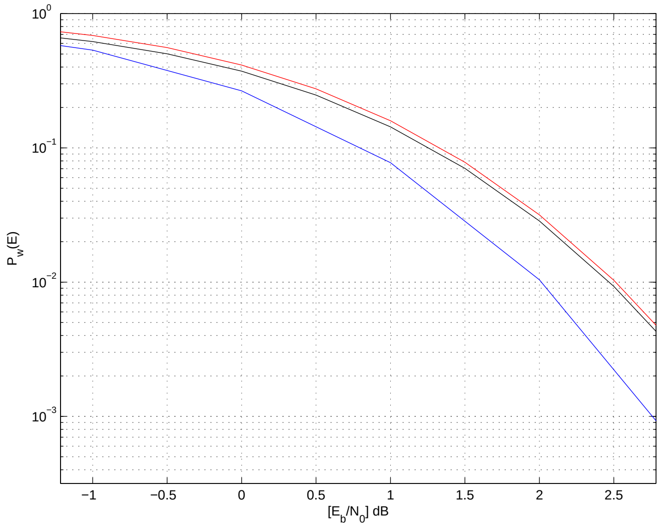

Fig. 4. Bounds on word error probability for BCH(63,30); from top to bottom: TSB, AHP with $w = 34$, and SLB.

to finding an optimum directed spanning graph for a vertex set. We address the above-mentioned problems by the addition of certain hyper-planes 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 addition of complexity.

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

It is also to be noted that, although it has been the tightest upper bound on the performance of fixed codes, the TSB fails to achieve the random coding exponent [11]. Typically, the random coding exponent $E(R)$ as a function of the code rate R can be studied in the three regions of $0 \leq R \leq R_{c1}$, $R_{c1} \leq R \leq R_{c2}$, and $R_{c2} \leq R \leq C$; referred to as the expurgated, straight-line, and sphere-packing bounds, respectively. C is the channel capacity and the two critical rate values R_{c1} and R_{c2} are the boundaries of the three regions [11]. It is possible to show that for codes with average spectrum, the exponent of the TSB coincides with the random coding exponent in expurgated and straight-line regions but not in the sphere-packing region [26]. TB is also known to be asymptotically incorrect while the SB for an average code coincides asymptotically with the random coding bound. Although the difference between the random coding exponent and the exponent of TSB for binary codes is quite small (specially for low rate codes), since the codewords of a binary code are not uniformly distributed on the Euclidean sphere, the conical regions do not form a dense sphere packing and therefore TSB will not be expected to be asymptotically tight. In the AHP bound, the underlying Gallager region is adopted from the TSB and the difference is primarily due to the use of a second-order Bonferroni inequality in place of the union bound. Thus, it is natural to not expect the AHP to be asymptotically tight, and in fact, our results are consistent with this observation offering no significant improvement for the error exponent. Though, It is worth noticing that the second version of Duman and Salehi (DS2) bound [12], [42], on the other hand, does give the correct random coding exponent [5], [29] (also relevant is the very recent works of Cohen et al. [4]).

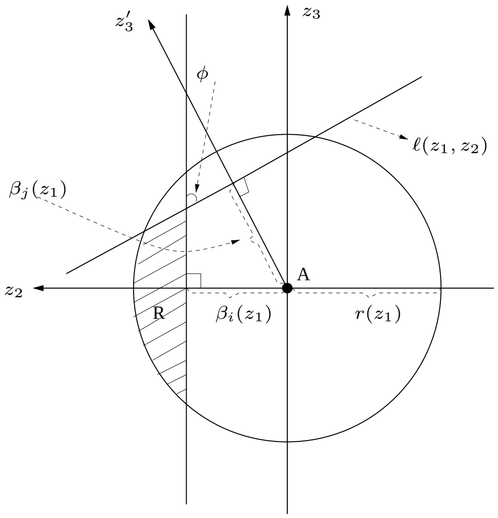

Fig. 5. Geometry for the probabilities $P_{z_1}(E_i, E_j^c, y \leq r^2(z_1))$ in the $z_2 - z_3$ plane.

Appendix A
Proof of (22) and (23)

Using the geometry of Fig. 5 in the plane of point A and noise vectors z_2 and z_3' defined in directions given in (21), we have

$$P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) = P_{z_1}\Big(\beta_i(z_1) \leq z_2, z_3' \leq \beta_j(z_1), y \leq r^2(z_1)\Big) = P(\mathbf{r} \in \mathrm{R}) \tag{40}$$

where $\mathbf{r}$ is the received vector from the channel and R is the crosshatched region shown in Fig. 5. (40) in terms of two orthogonal noise components z_2 and z_3 will be

$$P_{z_1}(E_i, E_j^c, y \leq r^2(z_1)) = P\Big(\beta_i(z_1) \leq z_2 \leq r(z_1), -r(z_1) \leq z_3 \leq \ell(z_1, z_2), y_2 \leq r^2(z_1) - z_2^2 - z_3^2\Big) \tag{41}$$

where $\ell(z_1, z_2)$ is the representation for the straight line perpendicular to z_3' as shown in Fig. 5. In above and (22), z_2 and z_3 are independent zero-mean Gaussian random variables with a variance of σ^2.

APPENDIX B
PROOF OF (28)

To study the correlation coefficient

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

we note that

$$\mathbf{s}_0 = \mathbf{m}(\mathbf{c}_0) = (\underbrace{-1 - 1 \cdots - 1}_{n}).$$

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 say that $\overrightarrow{v}$ belongs to the class $\{(\alpha)^d, (\gamma)^{n-d}\}$ containing $\left(\begin{array}{c} n \\ d \end{array}\right)$ such vectors, i.e.,

$$\overrightarrow{v} \in \{(\alpha)^d, (\gamma)^{n-d}\}. \tag{42}$$

Using the above class notation, $\mathbf{s}_i$ and $\mathbf{s}_j$ are also n-tuples with ± 1 components, where the number of ones in them is d_i and d_j respectively, i.e.,

$$\mathbf{s}_i \in \{(+1)^{d_i}, (-1)^{n-d_i}\}$$
$$\mathbf{s}_j \in \{(+1)^{d_j}, (-1)^{n-d_j}\}$$

and thus,

$$\overrightarrow{\mathbf{s}_0\mathbf{s}_i} \in \{(2)^{d_i}, (0)^{n-d_i}\}$$
$$\overrightarrow{\mathbf{s}_0\mathbf{s}_j} \in \{(2)^{d_j}, (0)^{n-d_j}\}$$

Also,

$$\overrightarrow{n_1} = \frac{\overrightarrow{\mathbf{s}_0\mathbf{o}}}{\|\overrightarrow{\mathbf{s}_0\mathbf{o}}\|}$$

Using (21) along with the definition of $\overrightarrow{n_1}$ and straightforward manipulations we have,

$$\overrightarrow{n_2} \in \frac{1}{\sqrt{\frac{d_i}{n}(n-d_i)}} \left\{(1-\frac{d_i}{n})^{d_i}, (-\frac{d_i}{n})^{n-d_i}\right\} \tag{43}$$

and

$$\overrightarrow{n_3} \in \frac{1}{\sqrt{\frac{d_j}{n}(n-d_j)}} \left\{(1-\frac{d_j}{n})^{d_j}, (-\frac{d_j}{n})^{n-d_j}\right\}. \tag{44}$$

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,

$$\rho_{max} = \begin{cases} \frac{d_j\left(n-d_i\right)}{\sqrt{d_i d_j (n-d_i)(n-d_j)}} = \sqrt{\frac{d_j\left(n-d_i\right)}{d_i\left(n-d_j\right)}} & \text{if} \quad d_i \geq d_j \\ \frac{d_i\left(n-d_j\right)}{\sqrt{d_i d_j (n-d_i)(n-d_j)}} = \sqrt{\frac{d_i\left(n-d_j\right)}{d_j\left(n-d_i\right)}} & \text{if} \quad d_i \leq d_j \end{cases} \tag{45}$$

and for minimum ρ, we have,

$$\rho_{min} = \begin{cases} -\sqrt{\frac{d_i d_j}{(n-d_i)(n-d_j)}} & \text{if} \quad d_i + d_j \leq n \\ -\sqrt{\frac{(n-d_i)(n-d_j)}{d_i d_j}} & \text{if} \quad d_i + d_j \geq n \end{cases}. \tag{46}$$

(45) and (46) together are equivalent to (28).

APPENDIX C: 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

$$\begin{aligned} P_{z_1}(E, y \leq r^2(z_1)) &\leq P_{z_1}(E, y \leq r^2(z_1), \mathbf{r} \in \Re_2) + P_{z_1}(y \leq r^2(z_1), \mathbf{r} \notin \Re_2) \\ &= P_{z_1}(E, y \leq r^2(z_1), \mathbf{r} \in \bigcap_{j=1}^{\eta} H_j^c) + P_{z_1}(y \leq r^2(z_1), \mathbf{r} \in \bigcup_{j=1}^{\eta} H_j) \end{aligned} \tag{47}$$

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

$$\begin{aligned} P_{z_1}(y \leq r^2(z_1), \mathbf{r} \in \bigcup_{j=1}^{\eta} H_j) &= P_{z_1}(\bigcup_{j=1}^{\eta} (\mathbf{r} \in H_j, y \leq r^2(z_1))) \\ &\leq \sum_{j=1}^{\eta} P_{z_1}(\mathbf{r} \in H_j, y \leq r^2(z_1)) \\ &= \eta P(\beta_w(z_1) \leq z_2 \leq r(z_1), y_1 \leq r^2(z_1) - z_2^2) \end{aligned} \tag{48}$$

with

$$\beta_w(z_1) = \frac{\sqrt{n} - z_1}{\sqrt{\frac{n}{w} - 1}}. \tag{49}$$

The last equality in (48) is simply due to the SB with a geometry similar to that of Fig. 1 and expressions in (10).

Also,

$$\begin{aligned} P_{z_1}(E, y \leq r^2(z_1), \mathbf{r} \in \bigcap_{j=1}^{\eta} H_j^c) &= P_{z_1}\Big(\bigcup_{k=1}^{M} E_k, y \leq r^2(z_1), \mathbf{r} \in \bigcap_{j=1}^{\eta} H_j^c\Big) \\ &= P_{z_1}\Big(\bigcup_{k=1}^{M} (E_k, y \leq r^2(z_1), \mathbf{r} \in \bigcap_{j=1}^{\eta} H_j^c)\Big) \\ &\leq P_{z_1}\Big(\bigcup_{k=1}^{M} (E_k, y \leq r^2(z_1), \mathbf{r} \in H_{j_{opt}^k}^c)\Big) \\ &\leq \sum_{k=1}^{M} P_{z_1}\Big(E_k, y \leq r^2(z_1), \mathbf{r} \in H_{j_{opt}^k}^c\Big) \end{aligned} \tag{50}$$

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

APPENDIX D
CALCULATION OF $\rho_{max,w}$ IN (37)

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 from the class

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

It is trivial to see that the maximum correlation will be

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

REFERENCES

[1] 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.

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

[3] 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.

[4] A. Cohen and N. Merhav, “Lower bounds on the error probability of block codes based on improvements on de Caen’s inequlaity,” *IEEE Trans. Inform. Theory*, vol. IT-52, no. 2, pp. 290-310, Feb. 2004.

[5] 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.

[6] S. Yousefi, “Bounds on the performance of maximum-likelihood decoded binary block codes in AWGN interference,” PhD dissertation, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, Sept. 2002.

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

[8] D. Divsalar, H. Jin, and R. J. McEliece, “ Coding theorems for ’Turbo-like’ codes,” *1998 Allerton Conference*, Sept. 23-25, 1998.

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

[10] 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.

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

[12] 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.

[13] —, “Performance bounds for Turbo-coded modulation systems,” *IEEE Trans. Commun.*, vol. COM-47, no. 4, pp. 511-521, Apr. 1999.

[14] —, “The union bound for Turbo-coded modulation systems over fading channels,” *IEEE Trans. Commun.*, vol. COM-47, no. 10, pp. 1495-1502, Oct. 1999.

[15] 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.

[16] —, “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.

[17] —, “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.

[18] 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.

[19] 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.

[20] 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.

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

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

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

[24] 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.

[25] —, “The error probability of M-ary PSK block-coded modulation schemes,” *IEEE Trans. Commun.*, vol. COM-44, no. 4, pp. 427-433, Apr. 1996.

[26] 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.

[27] 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.

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

[29] I. Sason and S. Shamai(Shitz), “Variations on the Gallager bounds, connections and applications,” *IEEE Trans. Inform. Theory*, vol. IT-48, no. 12, pp. 3029-3051, Dec. 2002.

[30] S. Yousefi and A. K. Khandani, “Generalized Tangential Sphere Bound on the ML Decoding Error Probability of Linear Binary Block Codes in AWGN Interference,” *IEEE Trans. Inform. Theory*, to be published.

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

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

[33] D. Hunter, “An upper bound for the probability of a union,” *J. Appl. Probab.*, vol. 13, pp. 597-603, 1976.

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

[35] J. B. Kruskal, “On the shortest spanning tree of a graph and the traveling salesman problem,” *Proc. of Amer. Math. Soc.*, no. 7, pp. 48-50, 1956.

[36] R. C. Prim, “Shortest connection networks and some generalizations,” *Bell Syst. Tech. Journ.*, no. 36 pp. 1389-1401, 1957.

[37] R. L. Graham and P. Hell, “On the history of the minimum spanning tree problem,” *Annals of the History of Computing*, vol. 7(1), pp. 43-57, 1985.

[38] K. H. Rosen, Discrete Mathemstics and its Applications. New York, NY: McGraw-Hill, 1995.

[39] 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.

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

[41] 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.

[42] T. M. Duman, “Turbo codes and Turbo-coded modulation systems: analysis and performance bounds,” PhD dissertation, Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, May 1998.

Shahram Yousefi received his B.Sc. in Electrical Engineering from University of Tehran, Iran, in 1996, ranking second in his graduating class. In Sept. 1997, he moved from industry to join the department of Electrical and Computer Engineering of University of Waterloo where he received his Ph.D. degree in electrical engineering in Sept. 2002. He then moved to Kingston, Ontario, Canada, where he is currently an Assistant Professor at the Department of Electrical and Computer Engineering of Queen’s University. Dr. Yousefi is the recipient of more than 20 awards and scholarships including the Natural Sciences and Engineering Research Council of Canada award, the Sandford Fleming Foundation Award, and the Golden Apple Award at Queen’s University. His research areas of interest are in the general areas of communication and Information theory, in particular, Channel coding/decoding, performance evaluation of codes, and the application of graphical representations in decoding.

Amir K. Khandani received his M.A.Sc. degree from University of Tehran and Ph.D. degree from McGill University, in 1985 and 1992, respectively. Following that, he worked for one year as a Research Associate at the INRS-Telecommunication, Montreal. In 1993, he joined the Department of Electrical and Computer Engineering of the University of Waterloo where he is currently working as a professor. Dr. Khandani is currently holding a NSERC-Nortel Networks senior industrial research chair in ”Advanced Telecommunications Technologies”. He is also acting as an associate editor for IEEE Transactions on Communications in the area of Coding and Communication Theory.