Extending Divsalar's Bound to Nonbinary Codes with two Dimensional Constellations

Dariush Divsalar
Jet Propulsion Laboratory.
California Institute of Technology
Pasadena, California, USA
Email: Dariush.Divsalar@jpl.nasa.gov

***Abstract*—Closed form upper bounds to error probabilities of nonbinary coded systems with two and higher dimensional constellation over the additive white Gaussian noise channel are derived. Computation of these bounds do not require integrations nor parameter optimizations. These bounds are the extension of bounds derived by Divsalar in 1999 for binary codes with binary modulations. The proposed bounds require only the knowledge of the average pairwise Euclidean distance enumerator of the code words when certain symmetry conditions do not hold. The bounds are tight for large block lengths. We also briefly discuss how to extend the pairwise Euclidean distance enumerators to ensembles of nonbinary protograph codes using the notion of frequency weight enumerators.**

I. Introduction

Modern codes, such as LDPC codes, have spurred a large amount of research on deriving tight bounds on error probability using maximum likelihood decoding [2]. These codes were first analyzed using the union bound, which cannot predict the performance above the cutoff rate. As the performance of modern codes is close to Shannon's capacity limit for large block sizes, the bounds that can predict the performance of modern codes for rates above the cutoff rate would be useful.

During the 1990s a tight bound by Poltyrev [13] and the geometry bound proposed by Dolinar, Ekroot, and Pollara [3] were among the tightest bounds. Duman and Salehi [5], [4] proposed bounds based on results of Gallager [9] which are referred to DS1 and DS2 bounding techniques. During the same time, Viterbi and Viterbi [17], developed a simple closed form bound. Most notable contribution since 2000 on tight bounds are due to Sason and Shamai [12], [6], [2]. Around the same time Divsalar [7], [11] proposed closed form bounds that are asymptotically tight. Later contribution by Yousefi and Khandani [14] extended these results. All these bounds assumed binary modulation with binary codes. A notable early contribution on bounds for nonbinary modulation is due to Herzberg and Poltyrev in the 1990's [8]. Recent paper by Hof, Sason, and Shamai [1] uses Gallager type bounds for nonbinary codes over nonbinary input discrete memoryless channels offers a significant contribution to bounding the error probability in the recent years. The bound in [7] also extended to q-ary orthogonal modulation [10], which only requires the symbol Hamming weight distribution of nonbinary codes.

In this paper, we obtain closed form upper bounds on the word error probability of nonbinary codes with two and higher dimensional constellation over the additive white Gaussian noise (AWGN) channel. The proposed bounds are an extension of Divsalar's bound [7] to nonbinary modulation with nonbinary codes. The bounds use the basic bounding technique and the Chernoff bounds proposed by Gallager [9]. The two dimensional constellation can be viewed as equal energy signals such as M-ary phase shift keying (MPSK) or un-equal energy signals such as M-ary quadrature amplitude modulation (MQAM). We obtain two separate bounds, one for equal energy signals and the other for un-equal energy signals. The derived bounds are asymptotically tight but not as tight as the bounds in [1] for finite length codes that were derived using integration and parameter optimizations. These tight bounds used versions of the Gallager bounds proposed by Sason and Shamai [2]. We also obtain a closed form expression for the minimum signal-to-noise ratio (SNR) threshold above which the bound is useful. Our bound can predict the performance close to the capacity limit but cannot achieve the capacity limit for certain range of code rates. This can be demonstrated for random codes as the code block length goes to infinity using the expression for the minimum threshold of the bound. It was shown in [7] that the minimum threshold for the tangential sphere bound is identical to the minimum threshold of the Divsalar bound for binary codes with binary modulations.

In Section II, we provide some motivation in writing this paper after 13 years. In Section III, we present channel models and bounds. In Section IV, we present the derivation of the closed form bound for equal energy signals such as MPSK signal constellation. In Section V, the bound is extended to un-equal energy signals such as MQAM. In Section VI, we briefly discuss pairwise weight enumerators and its extension to nonbinary protograph code ensembles since the derived bounds require pairwise Euclidean distance enumerators in general. A summary of results is presented in Section VII.

II. Motivation

Deriving closed form bounds in this paper is motivated by recent progress in ensemble enumerators for nonbinary LDPC codes both un-structured and structured ones, such as nonbinary unconstrained and graph cover protograph codes. In particular for the first time the notion of frequency weight enumerators and pairwise frequency weight enumerators were used in [15] that enables the computation of average Euclidean distance enumeration. Such enumerators in conjunction with tight ML bounds, and in particular closed form ML bounds, enable a comparison of maximum likelihood (ML) thresholds and iterative decoding thresholds for asymptotic cases.

III. Channel model and bounds

Consider a nonbinary code $\mathcal{C}$ with length n and M_c codewords over alphabet $\mathcal{A} = \{0, 1, \ldots, q-1\}$. Let the codewords be $\mathbf{c}_i \in \mathcal{C}$ for $i = 0, 1, 2, \ldots, M_c - 1$. Consider transmission of a κ-dimensional signal constellation $\mathcal{Q} = \{Q_0, Q_1, \ldots, Q_{q-1}\}$, where $Q_i \in \mathcal{Q}$ is a κ-dimensional channel symbol with average energy normalized to 1 representing a κ-dimensional constellation point. For equal energy signals such as κ-dimensional MPSK $\parallel Q_i \parallel^2 = 1$ for all i. Although our interest is in two dimensional MPSK and MQAM but we are considering κ-dimensional modulation in order to unify our final results with the results in [7] for BPSK modulation (i.e. κ=1). Assume there is a one to one mapping from $\mathcal{A}$ into the channel signal set $\mathcal{Q}$. Let codeword $\mathbf{c}_i \in \mathcal{C}$ map one to one to the sequence of channel symbols $\mathbf{x}_i \in \mathcal{X}$ (can be referred to channel codewords or simply codewords) of length n (κ-dimensional) channel symbols for $i = 0, 1, 2, \ldots, M_c - 1$. Each component of $\mathbf{x}_i$ takes values in $\mathcal{Q}$. Consider transmission of a codeword $\mathbf{x}$, over the memoryless AWGN channel, modeled in the form of $\mathbf{y} = \gamma \mathbf{x} + \mathbf{n}$ where $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{n}$ are κn-dimensional vectors; in particular, $\mathbf{x}$ denotes the transmitted signal vector, $\mathbf{y}$ the received vector, $\mathbf{n}$ a random noise vector whose components are independent Gaussian random variables with mean zero and variance 1 per dimension, and $\gamma = \sqrt{2R_c log_2 q(E_b/N_o)}$; E_b is the information bit energy; R_c is the code rate of the nonbinary code; and $N_o/2$ is the two-sided power spectral density of a white Gaussian noise process at the receiver. When the channel signal components have equal energy (MPSK) then $\parallel \mathbf{x} \parallel^2 = n$. With ML decoding, the word error probability when $\mathbf{x}$ is transmitted may depend on $\mathbf{x}$. Thus, the average word error probability can be written in the form

$$\bar{P}(e) = \frac{1}{M_c} \sum_{\mathbf{x}} \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \neq \mathbf{x}} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right], \tag{1}$$

where $\{\mathbf{x} \to \widehat{\mathbf{x}}\}$ denote the "pairwise error event." This is the probability that when $\mathbf{x}$ is transmitted the Euclidean distance (ML decoding metric) between the received vector $\mathbf{y}$ and $\widehat{\mathbf{x}}$ is smaller that the distance between $\mathbf{y}$ and $\mathbf{x}$, that is, $\{\mathbf{x} \to \widehat{\mathbf{x}}\} \triangleq \{\mathbf{y} \;:\; \parallel \mathbf{y} - \widehat{\mathbf{x}} \parallel < \parallel \mathbf{y} - \mathbf{x} \parallel\}$. Notice that the set of $\mathbf{y}$ such that $\{\mathbf{x} \to \widehat{\mathbf{x}}\}$ occurs is a half-space in $\mathbb{R}^{\kappa n}$, the locus of the points whose distance from $\mathbf{x}$ exceeds the distance from $\widehat{\mathbf{x}}$. A bound tighter than the union bound can be obtained as follows. Let d denote the generic Euclidean distance of $\widehat{\mathbf{x}}$ from $\mathbf{x}$. Partition the set of all codewords $\widehat{\mathbf{x}}$ (excluding $\mathbf{x}$) into equivalence classes of vectors with the same Euclidean distance value of $d > 0$. Denote these by $\mathcal{N}_{x,d}$, and write

$$\begin{aligned} P(e) &= \mathbb{P}\left[\bigcup_{d \in \mathcal{D}_x} \bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \\ &\leq \sum_{d \in \mathcal{D}_x} \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \{\mathbf{x} \to \widehat{\mathbf{x}}\}\right] \\ &= \sum_{d \in \mathcal{D}_x} \mathbb{P}[e_d | \mathbf{x}], \end{aligned} \tag{2}$$

where $\mathcal{D}_x$ is the set of distances (does not include $d = 0$) from $\mathbf{x}$ of the vectors in $\mathcal{X}$, and

$$e_d \triangleq \bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \{\mathbf{x} \to \widehat{\mathbf{x}}\} \tag{3}$$

is the event that, when $\mathbf{x}$ is transmitted, at least one other $\widehat{\mathbf{x}}$ at distance d is nearer to $\mathbf{y}$ than $\mathbf{x}$. We now compute an upper bound to $P(e)$ based on (2). To do this, we use a technique advocated by Gallager in [9] and express $\mathbb{P}[e_d]$ as

$$\mathbb{P}[e_d] = \mathbb{P}[e_d,\, \mathbf{y} \in \mathcal{R}] + \mathbb{P}[e_d,\, \mathbf{y} \notin \mathcal{R}], \tag{4}$$

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

$$\mathbb{P}[e_d,\, \mathbf{y} \notin \mathcal{R}] \leq \mathbb{P}[\mathbf{y} \notin \mathcal{R}]. \tag{5}$$

The region $\mathcal{R}$ should be chosen to minimize the overall word error rate. In practice, the selection of $\mathcal{R}$ should be guided by computational simplicity. By combining (5) and (4) we obtain

$$\mathbb{P}[e_d] \leq \mathbb{P}[e_d,\, \mathbf{y} \in \mathcal{R}] + \mathbb{P}[\mathbf{y} \notin \mathcal{R}]. \tag{6}$$

For equal energy constellations such as κ-dimensional MPSK we choose the following region, as was done in [7]. The bound is based on the choice for $\mathcal{R}$ of a κn-dimensional hypersphere centered at $\gamma\epsilon\mathbf{x}$ and with radius $\sqrt{\kappa n}R$,

$$\mathcal{R} = \{\mathbf{y} \mid \parallel \mathbf{y} - \epsilon\gamma\mathbf{x} \parallel^2 \leq \kappa n R^2\}. \tag{7}$$

The parameters ϵ and R will be selected so as to obtain the tightest possible bound.

A. Computation of $P[\mathbf{y} \notin \mathcal{R}]$

We compute the second term in (6). Define the RV W as

$$W = \parallel \mathbf{y} - \epsilon\gamma\mathbf{x} \parallel^2. \tag{8}$$

Using the Chernoff bound, we get

$$\mathbb{P}[\mathbf{y} \notin \mathcal{R}] = \mathbb{P}\left[W \geq \kappa n R^2\right] \leq e^{-\kappa s n R^2} \mathbb{E}[e^{sW}], \tag{9}$$

with $0 \leq s < \frac{1}{2}$. We observe that, under the assumption that $\mathbf{x}$ was transmitted, we have $\mathbf{y} = \gamma\mathbf{x} + \mathbf{n}$. Moreover, the components of $\mathbf{n}$ are independent RVs with mean zero and variance 1 per dimension. Thus, we obtain

$$\mathbb{E}[e^{sW}] = [\frac{1}{1-2s}]^{\frac{\kappa n}{2}} e^{\frac{s(1-\epsilon)^2}{1-2s}\gamma^2 \|\mathbf{x}\|^2}. \tag{10}$$

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

Define the RV Z as

$$Z = \parallel \mathbf{y} - \gamma\mathbf{x} \parallel^2 - \parallel \mathbf{y} - \gamma\widehat{\mathbf{x}} \parallel^2. \tag{11}$$

Then, we have

$$\begin{aligned} \mathbb{P}[e_d,\, \mathbf{y} \in \mathcal{R}] &= \mathbb{P}\left[\bigcup_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \{\mathbf{x} \to \widehat{\mathbf{x}}\},\, \mathbf{y} \in \mathcal{R}\right] \\ &\leq \sum_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \mathbb{P}\left[\{\mathbf{x} \to \widehat{\mathbf{x}}\},\, \mathbf{y} \in \mathcal{R}\right] \\ &= \sum_{\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}} \mathbb{P}\left[Z \geq 0,\; W \leq \kappa n R^2\right]. \end{aligned}$$

Using the Chernoff bound with parameters $t \geq 0$, $r \leq 0$, we get

$$\mathbb{P}[Z \geq 0,\, W \leq \kappa n R^2] \leq e^{-\kappa n r R^2} \mathbb{E}\left[e^{tZ+rW}\right], \tag{12}$$

and

$$\mathbb{E}\left[e^{tZ+rW}\right] = [\frac{1}{1-2r}]^{\frac{\kappa n}{2}} e^{\frac{\gamma^2}{1-2r} f(t,\, r,\, \epsilon)}, \tag{13}$$

where

$$\begin{aligned} f(t,\, r,\, \epsilon) &= (2t^2 - t + 2rt\epsilon) \| \, \widehat{\mathbf{x}} - \mathbf{x} \, \|^2 \\ &+ r(1-\epsilon)^2 \| \, \mathbf{x} \, \|^2 \\ &+ 2rt(1-\epsilon) \left[\| \, \widehat{\mathbf{x}} \, \|^2 - \| \, \mathbf{x} \, \|^2 \right]. \end{aligned} \tag{14}$$

The optimized parameter t^* is

$$t^* = \frac{1-2r\epsilon}{4} - \frac{r(1-\epsilon)}{2} \frac{\| \, \widehat{\mathbf{x}} \, \|^2 - \| \, \mathbf{x} \, \|^2}{\| \, \widehat{\mathbf{x}} - \mathbf{x} \, \|^2}, \tag{15}$$

provided $t \geq 0$ and $\| \, \widehat{\mathbf{x}} - \mathbf{x} \, \|^2 > 0$. We also note that the second derivative of the function in (14) with respect to t is positive. Note that we need the upper bound to be a function of Euclidean distance enumerator only. For non-equal energy signals such as MQAM this might be a problem. First, the difference in energies $\| \, \widehat{\mathbf{x}} \, \|^2 - \| \, \mathbf{x} \, \|^2$ can be positive, zero, or negative. Second, even if we use t^* given by (15) in (14), the function still depends on energies of codewords. So, to compute the bound, in addition to knowing the code distance enumerator, we need to know the energy of codewords. This result suggests that for unequal energy signals ϵ should be equal to 1. Thus, for unequal signal energy case, such as MQAM, the bound should be based on the choice for $\mathcal{R}$ of a κn-dimensional hypersphere centered at $\gamma \mathbf{x}$ and with radius $\sqrt{\kappa n} R$,

$$\mathcal{R} = \{\mathbf{y} \mid \| \, \mathbf{y} - \gamma \mathbf{x} \, \|^2 \leq \kappa n R^2\}. \tag{16}$$

In the next section we derive closed form bound for equal energy signals such as MPSK. For MQAM modulation even though the region is different we only need to use $\epsilon = 1$ in the results for MPSK.

IV. Closed form bound for MPSK modulation

For equal energy signals, we have $\| \, \mathbf{x} \, \|^2 = n$ for all codewords in $\mathcal{X}$. Thus $t^* = (1-2r\epsilon)/4$ and we obtain

$$f(t^*,\, r,\, \epsilon) = -\frac{(1-2r\epsilon)^2}{8} d^2 + r(1-\epsilon)^2 n. \tag{17}$$

Since $\widehat{\mathbf{x}} \in \mathcal{N}_{x,d}$ we have $\| \, \widehat{\mathbf{x}} - \mathbf{x} \, \|^2 = d^2$. Thus

$$\mathbb{E}\left[e^{t^* Z + rW}\right] = [\frac{1}{1-2r}]^{\frac{\kappa n}{2}} e^{\frac{\gamma^2}{1-2r}(-\frac{(1-2r\epsilon)^2}{8} d^2 + r(1-\epsilon)^2 n)}. \tag{18}$$

Equation (18) only depends on d^2, and (10) does not depend on the transmitted codeword. At this point we can perform averaging the word error probability over all transmitted codewords

$$\begin{aligned} \bar{P}(e) &\leq \frac{1}{M_c} \sum_{\mathbf{x}} \sum_{d \in \mathcal{D}_x} (|\mathcal{N}_{x,d}| e^{-\kappa r n R^2} \mathbb{E}[e^{t^* Z + rW}] \\ &+ e^{-\kappa s n R^2} \mathbb{E}[e^{sW}]). \end{aligned} \tag{19}$$

Define $\mathcal{D} = \bigcup_{\mathbf{x}} \mathcal{D}_x$ and let A_d be the averaged Euclidean distance enumerator over all transmitted codewords (average of $|\mathcal{N}_{x,d}|$). Then,

$$\bar{P}(e) \leq \sum_{d \in \mathcal{D}} (A_d e^{-\kappa r n R^2} \mathbb{E}[e^{t^* Z + rW}] + e^{-\kappa s n R^2} \mathbb{E}[e^{sW}]). \tag{20}$$

For each $d \in \mathcal{D}$, let $J_1 = A_d \mathbb{E}[e^{t^* Z + rW}]$ and $J_2 = \mathbb{E}[e^{sW}]$. Redefine parameters s, r, and ϵ as $2s = \frac{\beta-\rho}{1-\rho}$, $-2r = \frac{\beta-\rho}{\rho}$, and $\epsilon = \frac{\eta-\rho}{\beta-\rho}$, where $0 \leq \rho \leq \beta \leq 1$. Next, we minimize the upper bound in (20) with respect to the radius R for each $d \in \mathcal{D}$. The optimum value for R that minimizes the upper bound is

$$\sqrt{\kappa n} R^* = \sqrt{\frac{2\rho(1-\rho)}{(\beta-\rho)} \ln\left[\frac{\rho}{1-\rho} J_1^{-1} J_2\right]}. \tag{21}$$

Using this optimum R^* we obtain

$$\bar{P}(e) \leq c(\rho) \sum_{d \in \mathcal{D}} {J_1}^{\rho} {J_2}^{1-\rho}, \tag{22}$$

where $c(\rho) = (\frac{\rho}{1-\rho})^{1-\rho} + (\frac{1-\rho}{\rho})^{\rho} = 2^{H(\rho)} \leq 2$, and $H(\rho)$ is the binary entropy function. Since $\| \, \mathbf{x} \, \|^2 = n$, then for any $\mathbf{x}$ and $\mathbf{x}'$ in this coded (κ-dimensional) modulation scheme we have $d^2 = \| \, \mathbf{x} - \mathbf{x}' \, \|^2 \leq 4n$. We define the normalized squared Euclidean distance as $\delta^2 \triangleq \frac{d^2}{4n}$ where $0 \leq \delta^2 \leq 1$. After such normalization each $d \in \mathcal{D}$ is mapped one to one to $\delta \in \Delta$, where Δ represents the new set corresponding to the set $\mathcal{D}$. We normalize $\ln A_d$ as $r(\delta) = (\ln A_d)/n$. We also redefine the signal-to-noise ratio as $c = \gamma^2/2$. Substituting all new parameters we obtain

$$\bar{P}(e) \leq 2 \sum_{\delta \in \Delta} e^{-nE(c,\delta,\beta,\rho,\eta)}, \tag{23}$$

where

$$\begin{aligned} E(c,\delta,\beta,\rho,\eta) &= c\left[\frac{\eta^2\delta^2}{\beta} - \frac{(\eta-\beta)^2}{\beta(1-\beta)}\right] \\ &+ \frac{\kappa}{2}[\rho \ln\frac{\beta}{\rho} + (1-\rho)\ln\frac{1-\beta}{1-\rho}] - \rho r(\delta). \end{aligned} \tag{24}$$

Maximizing $E(c,\delta,\beta,\rho,\eta)$ with respect to η we get

$$\eta^* = \frac{\beta}{(1-\delta^2) + \beta\delta^2}. \tag{25}$$

The second derivative is negative. Using the optimum η^* we get

$$\begin{aligned} E(c,\delta,\beta,\rho) &= c\left[\frac{\beta\delta^2}{(1-\delta^2) + \beta\delta^2}\right] \\ &+ \frac{\kappa}{2}[\rho \ln\frac{\beta}{\rho} + (1-\rho)\ln\frac{1-\beta}{1-\rho}] - \rho r(\delta). \end{aligned} \tag{26}$$

Maximizing $E(c,\delta,\beta,\rho)$ with respect to ρ we get

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

The second derivative is negative. Using ρ^* we get

$$\begin{aligned} E(c,\delta,\beta) &= c\left[\frac{\beta\delta^2}{(1-\delta^2) + \beta\delta^2}\right] \\ &+ \frac{\kappa}{2}\ln\left[\beta + (1-\beta)e^{\frac{2}{\kappa}r(\delta)}\right] - r(\delta). \end{aligned} \tag{28}$$

The optimum β can be obtained as

$$\beta^* = \left[\sqrt{(1+\frac{c}{\kappa})^2 - 1 + \frac{c}{c_o(\delta)}} - (1+\frac{c}{\kappa})\right]\frac{1-\delta^2}{\delta^2}. \quad (29)$$

where

$$c_o(\delta) \triangleq (1-e^{-\frac{2}{\kappa}r(\delta)})\frac{1-\delta^2}{\frac{2}{\kappa}\delta^2}. \quad (30)$$

If the parameter $\beta^* \geq 1$, we set $\beta^* = 1$ for the union bound. The second derivative is negative. The bound is valid when $c > c_o(\delta)$. For κ=1, q=2, $d^2 = 4h$ (h is binary Hamming weight) we get the results in [7]. For $\kappa = q$, and $d^2 = 2h$ (h symbol Hamming weight of nonbinary code) we get the results for q-ary orthogonal modulation in [10]. For κ=2, the following bound

$$(\frac{E_b}{N_o})_{min} \leq \frac{1}{R_c \log_2 q}\max_{0<\delta<1}(1-e^{-r(\delta)})\frac{1-\delta^2}{\delta^2}. \quad (31)$$

can serve as the tightest closed form upper bound on the minimum signal-to-noise ratio threshold for ML decoding of coded MPSK.

V. Closed form bound for MQAM modulation

For unequal energy signals such as MQAM we argued that choosing the region given by (16) is more reasonable. This means that in the previous results for MPSK we can set $\epsilon = 1$, or equivalently $\eta = \beta$ in (24), In deriving the bound we don't need to have $\delta^2 \leq 1$ for MQAM case for this region. For κ-dimensional MQAM with average energy normalized to 1, we obtain

$$E(c,\delta,\beta,\rho) = c\beta\delta^2 + \frac{\kappa}{2}[\rho\ln\frac{\beta}{\rho} + (1-\rho)\ln\frac{1-\beta}{1-\rho}] - \rho r(\delta). \quad (32)$$

The optimum ρ is

$$\rho^* = \frac{\beta}{\beta + (1-\beta)e^{\frac{2}{\kappa}r(\delta)}}. \quad (33)$$

Using this optimum ρ^* we obtain

$$\begin{aligned} E(c,\delta,\beta) &= c\beta\delta^2 \\ &+ \frac{\kappa}{2}\ln\left[\beta + (1-\beta)e^{\frac{2}{\kappa}r(\delta)}\right] - r(\delta). \end{aligned} \quad (34)$$

The optimum β is

$$\beta^* = \frac{1}{\frac{2}{\kappa}\delta^2}(\frac{1}{c_o(\delta)} - \frac{1}{c}). \quad (35)$$

If the optimum $\beta^* \geq 1$ we set $\beta^* = 1$, which provides the union bound. The second derivative is negative, and

$$c_o(\delta) = (1-e^{-\frac{2}{\kappa}r(\delta)})\frac{1}{\frac{2}{\kappa}\delta^2}. \quad (36)$$

Then, we obtain

$$E(c,\delta) = \frac{\kappa}{2}[\frac{c}{c_o(\delta)} - 1 + \ln\frac{c_o(\delta)}{c}]. \quad (37)$$

The bound is valid when $c > c_o(\delta)$. For $\kappa = 2$ the bound

$$(\frac{E_b}{N_o})_{min} \leq \frac{1}{R_c \log_2 q}\max_{\delta>0}(1-e^{-r(\delta)})\frac{1}{\delta^2}. \quad (38)$$

can serve as the tightest closed form upper bound on the minimum signal-to-noise ratio threshold for ML decoding of coded MQAM.

VI. Pairwise weight enumerators for nonbinary protograph codes

A method for determining pairwise weight enumerators for a graph cover nonbinary protograph-based LDPC code ensemble has been proposed in [15]. In this section we extend the results in [15] to the ensembles of (unconstrained) nonbinary protograph based LDPC codes [16]. Given a base protograph G, the method is to create a modified protograph G^P from the original protograph G. Each variable (constraint) node in G^P can be viewed as a duplicated version of a variable (constraint) node in the original protograph G. The local adjacency is preserved among the replicated nodes. Each variable node in G^P takes a pair of elements of the field. This pair is interpreted as a new symbol in this set up. The frequency weight vector for a variable node i is defined as $\partial_i = [d_{i,0,0}, \ldots, d_{i,(q-1),(q-1)}]^T$ where $d_{i,\ell,\ell'}$ counts the number of occurrences (frequency) of the symbol (ℓ, ℓ') within N symbols. The definition of the frequency weight matrix for protograph G^P is the same as in [15]. The frequency weight matrix enumerator of the check node code $\mathcal{C}_{P,j}^N$ induced by the N copies of the constraint node $c_{P,j}$ with the associated scaling $\mathbf{s}_j$ was described in [15]. There, the frequency weight matrix enumerator of the code $\mathcal{C}_{P,j}^N$ was computed as $A^{\mathcal{C}_{P,j}^N}(\mathbf{d}_j)$ for given labels (non-zero elements of the GF(q)) attached to the edges of a check node. The elements of $\mathbf{d}_j$ comprise a subset of the elements of $\mathbf{d} = [\partial_1, \partial_2, ..., \partial_{n_v}]$, and this subset (corresponds to neighbors of $c_{P,j}$) is obtained from the edge connections in the mother protograph G^P. We have the following theorem as an extension of results to ensembles of nonbinary protograph based LDPC codes.

Theorem 1: The pairwise frequency weight matrix enumerator of the ensembles of nonbinary protograph based LDPC codes averaged over the entire ensemble is

$$\tilde{A}(\mathbf{d}) = \frac{\prod_{j=1}^{n_c} \bar{A}^{\mathcal{C}_{P,j}^N}(\mathbf{d}_j)}{\prod_{i=1}^{n_v} C(N;\, d_{i,0,0},\, d_{i,0,1},\, \ldots,\, d_{i,(q-1),(q-1)})^{t_i-1}}$$

where $\bar{A}^{\mathcal{C}_{P,j}^N}(\mathbf{d}_j)$ is the frequency weight matrix enumerator of the code $\mathcal{C}_{P,j}^N$ averaged over all possible associated scalings $\mathbf{s}_j$ and $C(N;\, n_0,\, n_1,\, \ldots, n_K)$ is multinomial coefficient. The proof is omitted for the lack of space. For any two dimensional modulation with constellation points $a(\ell)$, we can define the pairwise Euclidean distance [15] between two codewords as

$$d^2 = \sum_{i=1}^{n_v}\sum_{\ell=0}^{q-1}\sum_{\ell'=0}^{q-1} d_{i,\ell,\ell'}|a(\ell) - a(\ell')|^2. \quad (39)$$

Then, $A_d = \sum_{\mathbf{d}} \tilde{A}(\mathbf{d})$ is the pairwise weight enumerator where the sum ranges over all frequency weight matrix $\mathbf{d}$ that produce channel dependent parameter d^2. The computation of the enumerator can be simplified since d^2 is invariant to the order of (ℓ, ℓ') and further reduced when $\ell = \ell'$. Note that this result is particularly useful when the assumption on all-zeros transmission does not hold and can be used in our upper bounds for MPSK and MQAM modulations.

Example 1: Consider MPSK modulation where $M = q$. These closed form bounds predict the ML error performance beyond the channel cutoff rate and very close to the channel capacity as the block size goes to ∞. To see this we compute the minimum $\frac{E_b}{N_o}$ threshold that the bound can predict for a given throughput (rate) $R_c \log_2 q$. We already obtained this threshold for MPSK modulation as

$$(\frac{E_b}{N_o})_{th} = \frac{1}{R_c \log_2 M} \max_{\delta}(1 - e^{-r(\delta)}) \frac{1-\delta^2}{\delta^2}. \quad (40)$$

Consider random ensemble of non-binary codes of rate R_c and block size n over $GF(q)$. The frequency weigh enumerator can be shown to be

$$A(n_o, n_1, \cdots, n_{q-1}) = C(n, n_0, \cdots, n_{q-1}) q^{-n(1-R_c)} \quad (41)$$

where n_i for $i = 0, 1, \cdots, q-1$ represent the number of occurrences (frequency) of the nonbinary symbol i within the block length n. The Euclidean distance enumerator is then can be obtained as $A(d) = \sum_{\mathbf{n}} A(n_o, n_1, \cdots, n_{q-1})$ such that $\sum_{i=0}^{q-1} n_i d_i^2 = d^2$. For MPSK constellation ($q = M$) with unit radius $d_i^2 = 4\sin^2(\frac{\pi i}{M})$. The asymptotic Euclidean distance enumerator is $r(\delta) = \limsup_{n\to\infty} \frac{1}{n} \ln A_{\sqrt{4n\delta^2}}$ that is $r(\delta) = \max_{\mathbf{f}} H(f_0, f_1, \ldots, f_{M-1}) - (1-R_c)\ln M$ such that $\sum_{i=0}^{M-1} f_i \sin^2(\frac{\pi i}{M}) = \delta^2$, where H is the entropy function, $f_i = \frac{n_i}{n}$, and $\delta^2 = \frac{d^2}{4n}$. This can be computed as

$$r(\delta) = -\lambda\delta^2 + \ln \sum_{i=0}^{M-1} e^{\lambda \sin^2(\frac{\pi i}{M})} - (1-R_c)\ln M \quad (42)$$

such that λ satisfies

$$\sum_{i=0}^{M-1} \sin^2(\frac{\pi i}{M}) e^{\lambda \sin^2(\frac{\pi i}{M})} = \delta^2 \sum_{i=0}^{M-1} e^{\lambda \sin^2(\frac{\pi i}{M})} \quad (43)$$

Next we use this $r(\delta)$ in equation (40). The throughput (rate) versus the computed $(\frac{E_b}{N_o})_{th}$ threshold is plotted in Figure 1. Capacity of MPSK and the union bound threshold (cutoff rate) are also shown for comparison.

VII. Conclusion

Closed form upper bounds for error probabilities of nonbinary coded systems with two and higher dimensional MPSK and MQAM constellations over the additive white Gaussian noise channel are derived.

Acknowledgment

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. Research is supported in part by NSF grant CCF-1161822 (JPL Task Plan 82-17473).

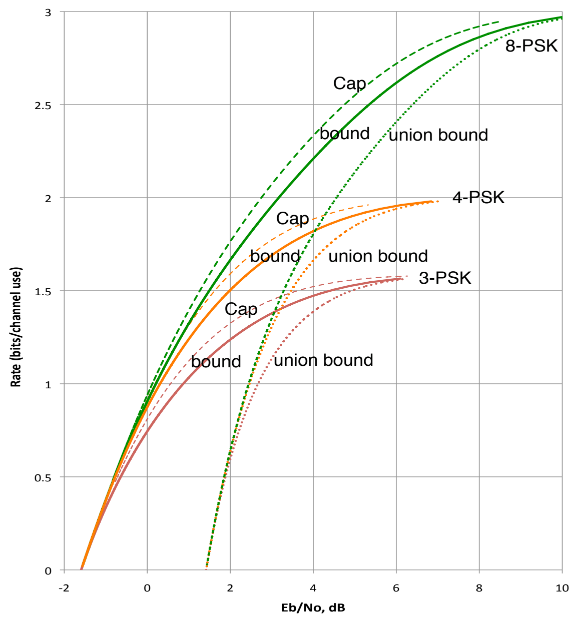

Fig. 1. Derived ML threshold (bound) for MPSK modulation using ensemble of random nonbinary codes is compared with capacity (Cap) and the union bound threshold (cutoff rate).

References

[1] Eran Hof, Igal Sason, and Shlomo Shamai (Shitz), Performance Bounds for Nonbinary Linear Block Codes Over Memoryless Symmetric Channels, *IEEE Trans. on Info. Theory*, Vol. 55, No. 3, March 2009.

[2] I. Sason and S. Shamai, Performance Analysis of Linear Codes under Maximum-likelihood Decoding: A Tutorial, Foundations and Trends in Communications and Information Theory, vol. 3, no. 1-2, pp. 1-222, June 2006.

[3] S. Dolinar, L. Ekroot, and F. Pollara,"Improved error probability bounds for block codes on the Gaussian channel, *IEEE ISIT* 1994.

[4] T. M. Duman, Turbo Codes and Turbo-Coded Modulation Systems: Analysis and Performance Bounds, Ph.D. dissertation, Elect. Comput. Eng. Dep., Northeastern University, May 1998.

[5] T. M. Duman and M. Salehi, New performance bounds for turbo codes, *IEEE Trans. on Comm.*, vol. 46, pp. 717-723, June 1998.

[6] S. Shamai and I. Sason, Variations on the Gallager bounds, connections and applications, *IEEE Trans. on Info. Theory*, Vol. 48, pp. 3029-3051, December 2002.

[7] D. Divsalar, A simple tight bound on error probability of block codes with application to turbo codes, the Telecommunications and Mission Operations (TMO) Progress Report 42139, JPL, pp. 1-35, November 15, 1999. [Online]. Available: http://tmo.jpl.nasa.gov/ progress report/42-139/139L.pdf.

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

[9] R. G. Gallager, Low-density parity-check codes, MIT 1963.

[10] Y. Kim, K. Cheun, and H. Lim, Performance of Weighted Nonbinary Repeat-Accumulate Codes over GF(q) with q-ary Orthogonal Modulation," *IEEE Trans. on Comm.*, May 2011.

[11] D. Divsalar and E. Biglieri, Upper bounds to error probabilities of coded systems beyond the cutoff rate, *IEEE Trans. on Comm.*, vol. 51, no. 12, pp. 2011-2018, December 2003.

[12] I. Sason and S. Shamai, Improved upper bounds on the ensemble performance of ML decoded low-density parity-check codes, *IEEE Comm. Letters*, vol. 4, no. 3, pp. 89-91, March 2000.

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

[14] Shahram Yousefi, and Amir K. Khandani,Generalized Tangential Sphere Bound on the ML Decoding Error Probability of Linear Binary Block Codes in AWGN Interference, *IEEE Trans. on Info. Theory*, Vol. 50, No. 11, November 2004.

[15] D. Divsalar and L. Dolecek,"Graph Cover Ensembles of Nonbinary Protograph LDPC Codes," *Proc. of IEEE ISIT* 2012.

[16] D. Divsalar and L. Dolecek,"Enumerators for Protograph-Based Ensembles of Nonbinary LDPC Codes,"*Proc. of IEEE ISIT* 2011.

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