Tight Performance Bounds for Permutation Invariant Binary Linear Block Codes over Symmetric Channels

Kostis Xenoulis, *Student Member, IEEE* and Nicholas Kalouptsidis

***Abstract*—Random coding performance bounds for L-list permutation invariant binary linear block codes transmitted over output symmetric channels are presented. Under list decoding, double and single exponential bounds are deduced by considering permutation ensembles of the above codes and exploiting the concavity of the double exponential function over the region of erroneous received vectors. The proposed technique specifies fixed list sizes L for specific codes under which the corresponding list decoding error probability approaches zero in a double exponential manner. The single exponential bound constitutes a generalization of Shulman-Feder bound and allows the treatment of codes with rates below the cutoff limit. Numerical examples of the new bounds for the specific category of codes are presented.**

***Index Terms*—Double exponential function, L-list permutation invariant codes, list decoding error probability, reliability function.**

I. Introduction

ERROR probability is a significant indicator of the transmission efficiency of coded information through a wide variety of communication channels. Closed form expressions for the calculation of the error probability are hard to establish in general, and thus tight analytical bounds are sought [1]. Classical treatments [2, Ch. 5] as well as modern approaches [3, Secs. 2-3] provide tight bounds mostly for random and structured families of codes (turbo codes [4], LDPC codes [5]), since the latter are treated more easily than specific codes. Thus the existence of at least one optimum code within these families is ascertained, but its respective characteristics remain unknown. On the other hand, random coding techniques, even though uncapable of finding optimum coding schemes, do provide bounds for information transmission rates up to the channel capacity and thus are preferable. Preliminary work towards combining random coding bounds for specific codes is reported in [6], where artificial ensembles of codes, invariant to error decoding probability, are constructed. The development of new tight bounding techniques is crucial to the design of specific codes with optimum characteristics which can achieve arbitrarily low error decoding probability with rates close to the channel's capacity.

In this paper, a new tight upper bound on the list decoding error probability of specific classes of codes over binary input symmetric output memoryless channels is derived. It is motivated by the effort to smooth out the influence of the union bound effect in the evaluation of the error probability. The proposed bound relies on the double exponential function rather than the standard exponential employed in the Chernoff bounding technique, as for example in [7]. The random coding argument is employed in the analysis through the set of all possible symbol positions permutations of the given code. Due to the symmetry of the channel and the specific structure of the treated codes, the list decoding error region is common to all codes of the ensemble. Furthermore, the relaxation of the new bound to a single exponential one, allows the derivation of a random coding bound for the previous category of codes with rates below the cutoff limit. This is in accordance to [8], where lower bounds on the list decoding error probability are derived for all code rates below capacity.

In the rest of this work, Section II provides a new upper bound on list decoding error probability and moreover introduces the notion of L-list permutation invariant codes. The random coding argument is applied to the previous class of codes and in relation with the new upper bound it leads, in Section III, to a tight bound on the L-list decoding error probability. Additionally, the extension of the generalized version of Shulman-Feder bound (SFB) [9, eq. (A17)], [10] is made feasible through the relaxation of the previous bound. Section IV provides numerical examples of the introduced bounds while discusses a way to lower bound tightly the reliability function of the considered channels. Finally, Section V concludes the present analysis and includes directions for further research.

II. Permutation Invariance and Error Probability Analysis

Consider the transmission of an arbitrary set of messages $\mathcal{M}$, with cardinality M, through a binary input, symmetric output discrete memoryless channel. Each message $m,\ 0 \leq m \leq M-1$, is encoded to an N-length codeword $\boldsymbol{c}_m$ that belongs to a given (N,R) binary, linear block code $\mathcal{C}$ and is transmitted through the channel. The codewords of $\mathcal{C}$ are denoted by $\boldsymbol{c}_0,\ldots,\boldsymbol{c}_{M-1}$, where $\boldsymbol{c}_0$ is the all-zero codeword, $\boldsymbol{c}_i \in \mathcal{I}^N, \mathcal{I}=\{0,1\}$, while $S_l,\ 0 \leq l \leq N$ with $S_0 = 1$ denotes the distance spectrum of $\mathcal{C}$. The minimum Hamming weight of the codewords in $\mathcal{C}$ is denoted by $d_{\min}$, while $wt_H(\boldsymbol{c}_i)$ is the Hamming weight of the binary vector $\boldsymbol{c}_i$. The transition probability measure of the channel, given the transmitted message m, is $\Pr(\boldsymbol{y}|\boldsymbol{c}_m)$, where $\mathbf{y}$ is the received vector at the output of the channel, also of length N. The output alphabet is $\mathcal{J}$ and hence $\boldsymbol{y} \in \mathcal{J}^N$. The error transition probability of the channel is $p/(q-1)$, where q is

The authors are with the Department of Informatics and Telecommunications, University of Athens, Greece (e-mail: {kxen,kalou}@di.uoa.gr)

TABLE I
COSET WEIGHT DISTRIBUTION MATRIX $\mathbf{\Gamma}$ OF THE REED-MULLER $\mathcal{R}(1,4)$ CODE.

Coset	Number of vectors of given weight																
Weight	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	1								30								1
1		1						15		15						1	
2			1				7		16		7				1		
3				1		3		12		12		3		1			
4					2		8		12		8		2				
4					4				24				4				
5						6		10		10		6					
6							16				16						

the cardinality of the channel's output alphabet $\mathcal{J}$. All vectors of length N considered in the analysis are column vectors.

Remark 1: Under the previous channel modeling setup, the proposed results apply either to a binary symmetric channel (BSC) with $q = 2$ or to the more general case of a binary-input, ternary-output channel with transition probability $P(0|1) = P(1|0) = p/(q-1)$ and erasure probability $P(e|1) = P(e|0) = p(q-2)/(q-1)$ (of course $P(0|0) = P(1|1) = 1-p$).

The translate $\boldsymbol{y} + \mathcal{C} = \{\boldsymbol{y} + \boldsymbol{c} : \boldsymbol{c} \in \mathcal{C}\}$ of a binary linear code $\mathcal{C}$ by any vector $\boldsymbol{y} \in \mathcal{J}^N$ is called a coset of the code, while a vector of smallest weight within the coset is called a coset leader. The weight distribution of a coset is defined in the same way as the distance spectrum of $\mathcal{C}$.

Definition 1: The coset weight distribution (or distance) matrix $\mathbf{\Gamma}$ of a binary linear block code $\mathcal{C}$ is defined as the $K \times (N+1)$ matrix having as rows all the K distinct coset weight distributions of $\mathcal{C}$. The first row of $\mathbf{\Gamma}$ is the distance spectrum $\{S_l\}_{l=0}^N$ of the code, while the other rows are sorted according to the weight of their coset leaders (coset weight). In the analysis following, matrix $\mathbf{\Gamma}$ is especially useful since it reveals the number of codewords a received vector $\boldsymbol{y}$ has within distance $wt_H(\boldsymbol{y})$. Namely, if a coset $\mathcal{C} + \boldsymbol{y}$ has weight distribution $\{\mathbf{\Gamma}_{\kappa,w}\}_{w=0}^N$ for some $\kappa \in [1, K]$, then $\boldsymbol{y}$ has $\mathbf{\Gamma}_{\kappa,w}$ codewords of $\mathcal{C}$ at exact distance w.

Letting list decoding be performed at the output of the channel with fixed list size L, the conditional error decoding probability of $\mathcal{C}$, given the transmission of message 0, satisfies

$$P_{e|0,\mathcal{C}}^{\mathcal{L}} = \sum_{\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}} \Pr(\boldsymbol{y}|\boldsymbol{c}_0, \mathcal{C}) \tag{1}$$

where

$$\boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}} = \{\boldsymbol{y} \in \mathcal{J}^N : \exists \{l_i\}_{i=1}^L, l_i \neq 0 : \Pr(\boldsymbol{y}|\boldsymbol{c}_{l_i}, \mathcal{C}) \geq \Pr(\boldsymbol{y}|\boldsymbol{c}_0, \mathcal{C}), \text{ for all } i \in [1, L]\}. \tag{2}$$

Since the channel is memoryless and output symmetric, $\boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$ is equivalently expressed as

$$\boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}} = \{\boldsymbol{y} \in \mathcal{J}^N : \exists \{l_i\}_{i=1}^L, l_i \neq 0 : wt_H(\boldsymbol{y} + \boldsymbol{c}_m) \leq wt_H(\boldsymbol{y}), \text{ for all } i \in [1, L]\}. \tag{3}$$

Moreover, if for $\lambda, \rho \geq 0$ we set

$$\Omega_L(\boldsymbol{y}, \mathcal{C}, \lambda, \rho) = \frac{L^\rho \Pr(\boldsymbol{y}|\boldsymbol{c}_0, \mathcal{C})^{\lambda\rho}}{\left(\sum_{m \neq 0} \Pr(\boldsymbol{y}|\boldsymbol{c}_m, \mathcal{C})^\lambda\right)^\rho} \tag{4}$$

then, due to the definition in (2), we have for $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$,

$$1 \leq e^{1 - e^{\Omega_L(\boldsymbol{y}, \mathcal{C}, \lambda, \rho) - 1}} \tag{5}$$

and thus the error decoding probability $P_{e|0,\mathcal{C}}^{\mathcal{L}}$ in (1) is upper bounded as

$$P_{e|0,\mathcal{C}}^{\mathcal{L}} \leq \sum_{\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}} \Pr(\boldsymbol{y}|\boldsymbol{c}_0, \mathcal{C}) e^{1 - e^{\Omega_L(\boldsymbol{y}, \mathcal{C}, \lambda, \rho) - 1}}. \tag{6}$$

In case $L = 1$, (6) consists an upper bound on the maximum likelihood (ML) error decoding probability. The current work is confined to the following L-list permutation invariant codes.

Definition 2: An (N, R) linear binary code $\mathcal{C}$ with coset weight distribution matrix $\mathbf{\Gamma}$ is L-list permutation invariant if both the following properties are satisfied:

$\mathfrak{L}_1$: there exists a $w_{opt} \geq \lceil d_{\min}/2 \rceil$ such that

$$L = \min_{\kappa \in [1,K], \mathbf{\Gamma}_{\kappa, w_{opt}} \neq 0} \mathbf{\Gamma}_{\kappa, w_{opt}} - 1 > 0$$

and

$$\max_{\kappa \in [1,K], w < w_{opt}} \mathbf{\Gamma}_{\kappa,w} < L + 1.$$

$\mathfrak{L}_2$: For all $\kappa \in [1, K]$, there exists a $w_\kappa^L > w_{opt}$ such that $\mathbf{\Gamma}_{\kappa, w_{opt}+1} = \ldots = \mathbf{\Gamma}_{\kappa, w_\kappa^L - 1} = 0$ and $\mathbf{\Gamma}_{\kappa, w_\kappa^L} \geq L + 1$.

If the previous conditions are met for $L = 1$, then we call the specific code ML permutation invariant.

Example 1: Consider the Reed-Muller $\mathcal{R}(1,4)$ code with the corresponding 8×17 coset weight distribution matrix $\mathbf{\Gamma}$, shown in Table I [11, Table 11.7]. The possible values of w_{opt} and L, under which conditions $\mathfrak{L}_1$, $\mathfrak{L}_2$ are met, are examined. For $w_{opt} = \lceil d_{\min}/2 \rceil = 4$, property $\mathfrak{L}_1$ is satisfied for $L = 1$. Moreover, for $w_1^1 = 8$, $w_2^1 = 7$, $w_3^1 = 6$, $w_4^1 = 5$, $w_5^1 = 6$, $w_6^1 = 8$, $w_7^1 = 5$, and $w_8^1 = 6$, all corresponding entries in $\mathbf{\Gamma}$ satisfy $\mathbf{\Gamma}_{\kappa, w_\kappa^1} \geq 2$. Consequently, property $\mathfrak{L}_2$ is also met by the specific code and $\mathcal{R}(1,4)$ is ML permutation invariant. If $w_{opt} = 5$, then $L = \min_{\kappa \in [1,8]} \mathbf{\Gamma}_{\kappa,5} - 1 = 2$ violates the second part of condition $\mathfrak{L}_1$, since $\max_{\kappa \in [1,8], w < 5} \mathbf{\Gamma}_{k,w} = 4 > L + 1 = 3$. For $w_{opt} = 6$, $L = \min_{\kappa \in [1,8]s} \mathbf{\Gamma}_{\kappa,6} - 1 = 6$ and

both conditions $\mathfrak{L}_1$, $\mathfrak{L}_2$ are satisfied. Thus $\mathcal{R}(1,4)$ is 6-list permutation invariant.

From an L-list permutation invariant code $\mathcal{C}$, we construct an ensemble of codes $\mathcal{E}$ by considering all possible symbol position permutation $N \times N$ matrices $\mathcal{P}$. A position permutation matrix $\mathcal{P}$ has a single 1 in every row and every column and is orthogonal, $\mathcal{P}\mathcal{P}^T = \mathcal{P}^T\mathcal{P} = I$, where I is the $N \times N$ identity matrix and $\mathcal{P}^T$ the transpose matrix of $\mathcal{P}$. Let for example

$$\mathcal{C} = \left\{\boldsymbol{c}_0^T = 000, \boldsymbol{c}_1^T = 011, \boldsymbol{c}_2^T = 100, \boldsymbol{c}_3^T = 111\right\}$$

and

$$\mathcal{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then, the permuted code is

$$\mathcal{P}\mathcal{C} = \left\{\mathcal{P}\boldsymbol{c}_0^T = 000, \mathcal{P}\boldsymbol{c}_1^T = 110, \mathcal{P}\boldsymbol{c}_2^T = 001, \mathcal{P}\boldsymbol{c}_3^T = 111\right\}$$

and $\mathcal{P}$ corresponds to the symbol position permutation $\Pi = \{2, 3, 1\}$. The lemma provided below is crucial in the derivation of the new tight bound on the list decoding error probability.

Lemma 1: For an (N, R) binary linear block code $\mathcal{C}$ that is L-list permutation invariant, all codes in the permuted ensemble $\mathcal{E}$ have the same error decoding region $\boldsymbol{Y}_0^{\mathcal{L}}$.

Proof: In the original L-list permutation invariant code $\mathcal{C}$, assume the received vector is $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$ where the weight distribution of the unique coset $\mathcal{C} + \boldsymbol{y}$ is $\{\boldsymbol{\Gamma}_{\nu,w}\}_{w=0}^N$. Due to the restriction imposed by condition $\mathfrak{L}_1$, it must be that $wt_H(\boldsymbol{y}) \geq w_{opt}$. Indeed, if $wt_H(\boldsymbol{y}) < w_{opt}$, then $\boldsymbol{y}$ will have at most $L-1$ codewords, not including the all zeros codeword $\boldsymbol{c}_0$, within Hamming distance $wt_H(\boldsymbol{y})$ and thus will not cause a list decoding error.

Let such a received vector $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$, with $wt_H(\boldsymbol{y}) = w_{opt}$. Then $\boldsymbol{y}$ has $\boldsymbol{\Gamma}_{\nu,w_{opt}} - 1 \geq L$ codeword(s), different from $\boldsymbol{c}_0$, at Hamming distance w_{opt}. Let, for any permutation matrix P, the corresponding permuted vector $P^T\boldsymbol{y}$ of $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$. Vector $P^T\boldsymbol{y}$ is also of weight w_{opt} while the coset weight distribution of $\mathcal{C} + P^T\boldsymbol{y}$ is denoted by $\{\boldsymbol{\Gamma}_{\xi,w}\}_{w=0}^N$, where ξ does not necessarily equal ν. Since L is selected as the minimum non-zero term of the w_{opt} column of $\boldsymbol{\Gamma}$ minus 1 and $\boldsymbol{\Gamma}_{\kappa,w} \leq L$ for all $\kappa \in [1, K], w < w_{opt}$ (property $\mathfrak{L}_1$), $P^T\boldsymbol{y}$ also has L codewords $\{\boldsymbol{c}_{l_i}\}_{i=1}^L$ of $\mathcal{C}$, apart from the all-zeros codeword $\boldsymbol{c}_0$, at distance $wt_H(\boldsymbol{y}) = w_{opt}$, i.e. $wt_H\left(P^T\boldsymbol{y} + \boldsymbol{c}_{l_i}\right) = wt_H(\boldsymbol{y}), i \in [1, L]$. Thus $P^T\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$. Moreover it is noted that

$$wt_H\left(P^T\boldsymbol{y} + \boldsymbol{c}_{l_i}\right) = wt_H\left(P\left(P^T\boldsymbol{y} + \boldsymbol{c}_{l_i}\right)\right) = wt_H\left(\boldsymbol{y} + P\boldsymbol{c}_{l_i}\right) = wt_H(\boldsymbol{y}). \tag{7}$$

The innermost equality in the right hand side of (7) indicates that $\boldsymbol{y}$ also has L codewords, apart from $\boldsymbol{c}_0$, of the permuted code $\mathcal{P}\mathcal{C}$ at distance $wt_H(\boldsymbol{y})$. Hence, $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$ causes a list decoding error in $\mathcal{P}\mathcal{C}$.

Let $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$ with weight $wt_H(\boldsymbol{y}) = w_\nu^{\mathcal{L}}$, for some $\nu \in [1, K]$, and the weight distribution of the corresponding coset $\mathcal{C} + \boldsymbol{y}$, $\{\boldsymbol{\Gamma}_{\nu,w}\}_{w=0}^N$. For any permutation matrix P, the weight distribution of coset $\mathcal{C}+P^T\boldsymbol{y}$ is denoted by $\{\boldsymbol{\Gamma}_{\xi,w}\}_{w=0}^N$, where ξ is not necessarily equal to ν. Then, due to property $\mathfrak{L}_2$, $P^T\boldsymbol{y}$ has $\boldsymbol{\Gamma}_{\xi,w_\xi^{\mathcal{L}}} - 1 \geq L$ codewords of $\mathcal{C}$, apart from $\boldsymbol{c}_0$, at Hamming distance $w_\xi^{\mathcal{L}} \leq w_\nu^{\mathcal{L}}$. Equality holds in the latter relation in case cosets $\mathcal{C} + \boldsymbol{y}$ and $\mathcal{C} + P^T\boldsymbol{y}$ have the same weight distribution. Consequently, $P^T\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$, while the weight transformation (7) guarantees that $\boldsymbol{y}$ also causes a list decoding error in the permuted code $\mathcal{P}\mathcal{C}$. Finally, in all other cases where $\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$ and $wt_H(\boldsymbol{y}) > w_\kappa^L$ for all $\kappa \in [1, K]$, $P^T\boldsymbol{y}$ always has $\min_{\kappa\in[1,K]} \boldsymbol{\Gamma}_{\kappa,w_\kappa^L} - 1 \geq L$ codewords of $\mathcal{C}$ within distance $wt_H(\boldsymbol{y})$ and thus $P^T\boldsymbol{y} \in \boldsymbol{Y}_{0,\mathcal{C}}^{\mathcal{L}}$. ■

In relation with the proof of Lemma 1, the proper selection of fixed list size L, under property $\mathfrak{L}_1$, guarantees that permutations of erroneous received vectors $\boldsymbol{y}$ of minimum weight are also erroneous. On the other hand, property $\mathfrak{L}_2$ assures that all received vectors of weight $wt_H(\boldsymbol{y}) > w_{opt}$ will have within distance $wt_H(\boldsymbol{y})$ at least L codewords of $\mathcal{C}$.

III. New Tight Upper Bounds

Due to the channel symmetry, the average list decoding error probability $P_e^{\mathcal{L}}$ of any code $\mathcal{C}$, over all messages in $\mathcal{M}$, equals $P_{e|0,\mathcal{C}}^{\mathcal{L}}$ [12, Appendix C]. Thus, any bound on $P_{e|0,\mathcal{C}}^{\mathcal{L}}$ is also a bound on $P_e^{\mathcal{L}}$. Moreover, for a L-list permutation invariant code $\mathcal{C}$, in all codes of the ensemble $\mathcal{E}$ constructed in Section II, message 0 is encoded into the all-zeros vector $\boldsymbol{c}_0$. Thus, $\Pr(\boldsymbol{y}|\boldsymbol{c}_0, \mathcal{C}) = \Pr(\boldsymbol{y}|\boldsymbol{c}_0)$. Consequently, if we take the average over $\mathcal{E}$ on both sides of (6), then due the error decoding region invariance property stated in Lemma 1, we have

$$P_{e|0}^{\mathcal{L}} \leq \sum_{\boldsymbol{y}\in\boldsymbol{Y}_0^{\mathcal{L}}} \Pr(\boldsymbol{y}|\boldsymbol{c}_0) E\left[e^{1-e^{\Omega_L(\boldsymbol{y},\mathcal{C},\lambda,\rho)-1}}\right]. \tag{8}$$

Note that the function $\exp(1 - \exp(x - 1))$ is concave for $0 \leq x \leq 1$ since

$$\frac{d^2 e^{1-e^{x-1}}}{dx^2} = e^{-1-e^{-1+x}+x}\left(-e + e^x\right) \leq 0, \text{ for } 0 \leq x \leq 1.$$

Moreover, for any $\boldsymbol{y} \in \boldsymbol{Y}_0^{\mathcal{L}}$, $\Omega_L(\boldsymbol{y}, \mathcal{C}, \lambda, \rho) \leq 1$. Therefore, application of Jensen's inequality to the right hand side of (8) gives

$$P_e^{\mathcal{L}} \leq \sum_{\boldsymbol{y}\in\boldsymbol{Y}_0^{\mathcal{L}}} \Pr(\boldsymbol{y}|\boldsymbol{c}_0) e^{1-e^{E[\Omega_L(\boldsymbol{y},\mathcal{C},\lambda,\rho)]-1}}. \tag{9}$$

The following technical lemma is useful in the derivation of a closed form upper bound on $P_e^{\mathcal{L}}$. The proof is provided in Appendix A.

Lemma 2: The mean value of the double exponent in (9) is lower bounded for all $\rho' \geq 0$ by

$$E\left[\Omega_L(\boldsymbol{y}, \mathcal{C}, \lambda, \rho)\right] \geq L^{\rho' P} \Pr(\boldsymbol{y}|\boldsymbol{c}_0)^{\frac{\rho'}{1+\rho'}} \left[(M-1)^{\rho' P} \cdot \left(\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)^{\rho'\frac{P}{Q}} \left(\sum_{\boldsymbol{x}\in\mathcal{I}^N} 2^{-N} \Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho'}}\right)^{\rho'}\right]^{-1} \tag{10}$$

where

$$b_l = \frac{\binom{N}{l}}{2^N}, \quad v_l = \frac{S_l}{M-1}, \quad 0 \le l \le N,$$
$$\frac{1}{P} + \frac{1}{Q} = 1, \quad P, Q \ge 1. \tag{11}$$

Combining Lemma 2 with (9) and passing from $\boldsymbol{Y}_0^{\mathcal{L}}$ to the set of all received vectors $\boldsymbol{Y}$, we get

$$P_e^{\mathcal{L}} \le \sum_{\boldsymbol{y} \in \mathcal{J}^N} \Pr(\boldsymbol{y}|\boldsymbol{c}_0) \exp\left(1 - \exp\left(\frac{2^{N\rho'} L^{\rho' P}}{(M-1)^{\rho' P}} \cdot \right.\right.$$
$$\left.\left. \frac{\Pr(\boldsymbol{y}|\boldsymbol{c}_0)^{\frac{\rho'}{1+\rho'}}}{\left(\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)^{\rho' \frac{P}{Q}} \left(\sum_{\boldsymbol{x} \in \mathcal{I}^N} \Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho'}}\right)^{\rho'}} - 1\right)\right). \tag{12}$$

The previous analysis introduces a new double exponential upper bound on the average L-list decoding error probability of an (N, R) binary, linear block code $\mathcal{C}$. The proof of the following theorem is provided in Appendix B.

Theorem 1: Consider an (N, R) binary linear block code $\mathcal{C}$ which is L-list permutation invariant with distance spectrum S_l, $0 \le l \le N$ and coset weight distribution matrix $\boldsymbol{\Gamma}$. $\mathcal{C}$ is utilized in the transmission of an arbitrary set of messages $\mathcal{M}$, with cardinality $M = 2^{NR}$, over a binary input, symmetric output discrete memoryless channel. If $p/(q-1)$ is the error transition probability of the channel, then the average L-list decoding error probability, over all messages in $\mathcal{M}$, $P_e^{\mathcal{L}}$ of $\mathcal{C}$ is upper bounded for all $\rho' \ge 0$ as

$$P_e^{\mathcal{L}} \le \sum_{h=0}^N \binom{N}{h} (1-p)^{N-h} \left(\frac{p}{q-1}\right)^h \sum_{k=\delta(q-2)h}^h \binom{h}{k} \cdot$$
$$(q-2)^{h-k} \exp\left(1 - \exp\left(\frac{(M-1)^{-\rho' P} 2^{N\rho'} L^{\rho' P}}{\left(\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)^{\rho' \frac{P}{Q}}} \cdot \right.\right.$$
$$\left.\left. \frac{\left((1-p)^{N-h}\left(\frac{p}{q-1}\right)^h\right)^{\frac{\rho'}{1+\rho'}}}{\mathcal{K}_1(p,q,\rho')^{N-h+k} \mathcal{K}_2(p,q,\rho')^{h-k}} - 1\right)\right) \tag{13}$$

where

$$\mathcal{K}_1(p,q,\rho') = \left((1-p)^{\frac{1}{1+\rho'}} + \left(\frac{p}{q-1}\right)^{\frac{1}{1+\rho'}}\right)^{\rho'},$$
$$\mathcal{K}_2(p,q,\rho') = \left(\frac{p}{q-1}\right)^{\frac{\rho'}{1+\rho'}}, \quad \frac{1}{P} + \frac{1}{Q} = 1, \quad P, Q \ge 1$$
$$\text{and} \quad \delta(q-2) = \begin{cases} 1, & q = 2 \\ 0, & \text{otherwise} \end{cases}. \tag{14}$$

We note that the upper bound of Theorem 1 fails to reproduce the random coding exponent for an L-list permutation invariant code $\mathcal{C}$, as in [6, Th. 1]. Additionally, it does not admit a closed form expression for continuous output channel. Nevertheless, since

$$e^{1-e^{x-1}} \le \frac{1}{x}, \quad x > 0 \tag{15}$$

the upper bound in (13) is tighter than the generalized version of Shulman-Feder bound in [9, eq. (A17)], [12, Cor. 8]. Moreover, for L-list permutation invariant codes, application of (15) in (13) provides a new version of the generalized SFB, which nicely complements the one presented in [9, eq. (A17)].

Theorem 2: Under the assumptions of Theorem 1, the following bound holds for $q \ge 2$ and all $\rho' \ge 0$

$$P_e^{\mathcal{L}} \le e^{-N\left(E_o(\rho') - \rho'\left(P(R - \frac{1}{N}\ln L) + \frac{P}{Q}\frac{1}{N}\ln\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)\right)}$$
$$\text{where } E_o(\rho') = -\ln\left(\left(\frac{1}{2}\right)^{\rho'}\left((1-p)^{\frac{1}{1+\rho'}} + \right.\right.$$
$$\left.\left.\left(\frac{p}{q-1}\right)^{\frac{1}{1+\rho'}}\right)^{1+\rho'} + (q-2)\frac{p}{q-1}\right) \tag{16}$$
$$\text{and } \frac{1}{P} + \frac{1}{Q} = 1, \quad P, Q \ge 1.$$

Proof: For $q > 2$, application of (15) in (13) and grouping of the terms $(1-p)^{N-h}$ and $(p/(q-1))^h$ yield

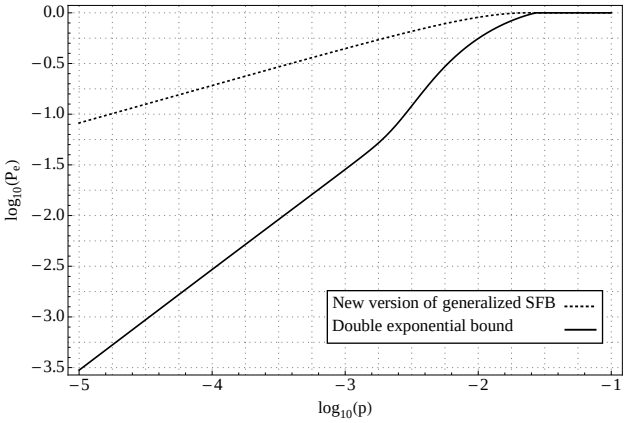
$$P_e^{\mathcal{L}} \le \left(\frac{M-1}{L}\right)^{\rho' P} \left(\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)^{\rho' \frac{P}{Q}} \left(\frac{1}{2}\right)^{\rho' N} \cdot$$
$$\sum_{h=0}^N \left[(1-p)^{\frac{1}{1+\rho'}}\right]^{N-h} \left[\left(\frac{p}{q-1}\right)^{\frac{1}{1+\rho'}}\right]^h \frac{N!}{(N-h)!h!} \cdot$$
$$\mathcal{K}_1(p,q,\rho')^{N-h} \sum_{k=0}^h \frac{h!}{k!(h-k)!} \mathcal{K}_1(p,q,\rho')^k \cdot$$
$$\left((q-2)\mathcal{K}_2(p,q,\rho')\right)^{h-k}. \tag{17}$$

Moreover due to the binomial expansion formula it holds

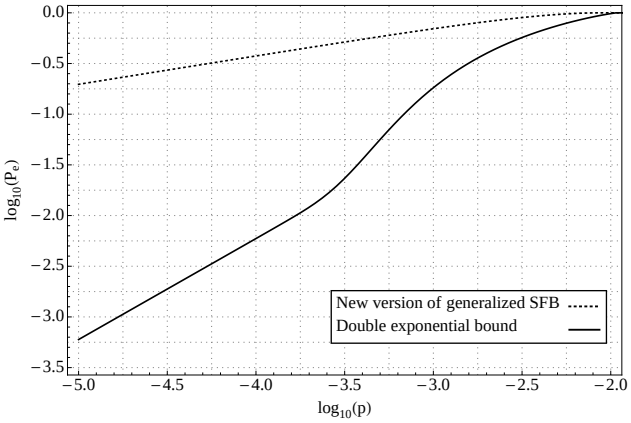
$$\sum_{k=0}^h \frac{h!}{k!(h-k)!} \mathcal{K}_1(p,q,\rho')^k \left((q-2)\mathcal{K}_2(p,q,\rho')\right)^{h-k} =$$
$$\left(\mathcal{K}_1(p,q,\rho') + (q-2)\mathcal{K}_2(p,q,\rho')\right)^h \tag{18}$$

so that replacing (18) in (17) and collecting the terms raised to the h and $(N-h)$-th powers respectively, we have

$$P_e^{\mathcal{L}} \le \left(\frac{M-1}{L}\right)^{\rho' P} \left(\sum_{l=0}^N b_l \left(\frac{v_l}{b_l}\right)^Q\right)^{\rho' \frac{P}{Q}} \left(\frac{1}{2}\right)^{\rho' N} \cdot$$
$$\sum_{h=0}^N \frac{N!}{(N-h)!h!} \left[(1-p)^{\frac{1}{1+\rho'}} \mathcal{K}_1(p,q,\rho')\right]^{N-h} \cdot$$
$$\left[\left(\frac{p}{q-1}\right)^{\frac{1}{1+\rho'}} \mathcal{K}_1(p,q,\rho') + \left(\frac{p}{q-1}\right)^{\frac{1}{1+\rho'}} \cdot \right.$$
$$\left. (q-2)\mathcal{K}_2(p,q,\rho')\right]^h. \tag{19}$$



(a) $\mathcal{R}(1,4)$ with $L=6$.



(b) $\mathcal{R}(2,5)$ with $L=5$.

Fig. 1. Comparison between the double exponential upper bound (21) and the new version (22) of the generalized SFB for the transmission of the L-list permutation invariant codes over a discrete memoryless BSC with transition probability p under list decoding with fixed list size L.

Repeating the above arguments to the sum in the right hand side of (19) and due to the definitions of $\mathcal{K}_1(p,q,\rho')$ and $\mathcal{K}_2(p,q,\rho')$ in (14), (16) is satisfied for $q > 2$. Similar arguments validate (16) for $q=2$. ∎

We note here the upper bound of Theorem 2 can result by direct application of (15) to (12). Furthermore, it is noted that the bounds of Theorems 1 and 2 cannot result as special cases of the bounding technique that leads to the DS2 bound [9, eq. (8)], since in the latter no proper selection of the unnormalized tilting measure can lead to the upper bound (12).

IV. Applications and Discussion

In order to simplify the calculation of the upper bounds of Theorems 1 and 2 and afford coding exponents resembling the random coding one, we proceed by removing the parameters P, Q in the respective theorems. Specifically, we note that

$$\frac{1}{N}\ln\left(\sum_{l=0}^{N} b_l\left(\frac{v_l}{b_l}\right)^{Q}\right)^{\frac{1}{Q}} \leq \frac{1}{N}\ln\max_{0\leq l\leq N}\left(\frac{v_l}{b_l}\right) = \mathcal{B}(\mathcal{C}). \tag{20}$$

Then, if we replace in (13) and (16) the upper bound (20) and moreover allow $P=1$ and $Q\to\infty$, then we have respectively for $\rho'\geq 0$

$$P_e^{\mathcal{L}} \leq \sum_{h=0}^{N}\binom{N}{h}(1-p)^{N-h}p^h \exp\left(1-\exp\left(\frac{L^{\rho'}}{(M-1)^{\rho'}}\cdot \frac{\left((1-p)^{N-h}p^h\right)^{\frac{\rho'}{1+\rho'}}}{e^{N\rho'\mathcal{B}(\mathcal{C})}\left(\frac{1}{2}\right)^{\rho' N}\mathcal{K}_1(p,q,\rho')^N}-1\right)\right) \tag{21}$$

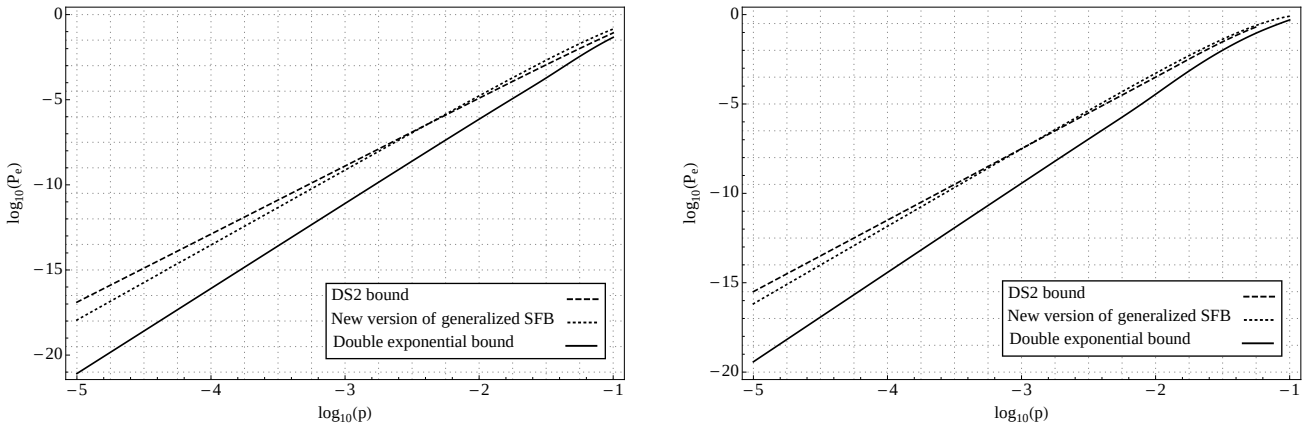
and

$$P_e^{\mathcal{L}} \leq e^{-N\left(E_o(\rho')-\rho'\left(R-\frac{1}{N}\ln L+\mathcal{B}(\mathcal{C})\right)\right)}. \tag{22}$$

The nature of the above bounds under list decoding with fixed list size L is illuminated by the following examples. Consider the transmission of an arbitrary set of messages $\mathcal{M}$ over a discrete memoryless binary symmetric channel (BSC) with error transition probability p. The 6-list permutation invariant $\mathcal{R}(1,4)$ code, presented in Table I is examined first. The weight distribution of the specific code, illustrated in the first line of the matrix in Table I, facilitates the calculation of the term in the right hand side of (20). Specifically, for $\mathcal{B}(\mathcal{R}(1,4))=0.478523$, the double exponential upper bound (21) and the new version (22) of the generalized Shulman-Feder bound are respectively optimized for $\rho'\geq 0$. Results are depicted in Fig. 1(a), where the tightness of (21) over (22) is illustrated. For the aforementioned transmission procedure, we consider next the Reed-Muller $\mathcal{R}(2,5)$ code, where the corresponding coset weight distribution is depicted in [11, Table 11.4]. The specific code is not ML permutation invariant since it does not satisfy property $\mathfrak{L}_1$. On the other hand, $L=5$ is the minimum value under which conditions $\mathfrak{L}_1, \mathfrak{L}_2$ are met. For the previous valid fixed list size, the optimized versions of the upper bounds (21) and (22) respectively are depicted in Fig. 1(b) for $\mathcal{B}(\mathcal{R}(2,5))=0.346574$. Furthermore, in both Figs. 1(a), 1(b), the double exponential upper bound curves resemble that of the double exponential function $\exp(1-\exp(x-1))$.

Finally, we perform a numerical comparison among the double exponential bound (13) of Theorem 1, the new version (16) of the generalized SFB of Theorem 2 and the DS2 bound [9, eq. (8)]. Again, we consider the transmission of L-list permutation invariant codes over a discrete memoryless binary symmetric channel (BSC) with error transition probability p, performing list decoding with fixed list size L. Under the previous setup, the DS2 bound is expressed for $0\leq\lambda$ and $0\leq\rho\leq 1$ by

$$P_{e|0} \leq \left(\frac{1}{L}\right)^{\rho}\left(\sum_{y}\Pr(y|0)g(y)\right)^{N(1-\rho)}\cdot \left(\sum_{d=d_{\min}}^{N} S_d\left(\sum_{y}\Pr(y|0)g(y)^{1-\frac{1}{\rho}}\right)^{N-d}\right.\cdot$$



(a) $\mathcal{R}(1,4)$ with $L=6$. (b) $\mathcal{R}(2,5)$ with $L=5$.

Fig. 2. Comparison among the double exponential upper bound (13), the new version (16) of the generalized SFB and the DS2 bound (23) for the transmission of L-list permutation invariant codes over a discrete memoryless BSC with transition probability p under list decoding with fixed list size L.

$$\left(\sum_{y}\Pr(y|0)^{1-\lambda}\Pr(y|1)^{\lambda}g(y)^{1-\frac{1}{\rho}}\right)^{d}\Bigg)^{\rho}. \qquad (23)$$

The optimized tilting measure $g(y)$ in (23) is found by numerical optimization for each value of the error transition probability p. Numerical results are presented in Figs. 2(a) and 2(b) for the 6-list permutation invariant $\mathcal{R}(1,4)$ code and the 5-list permutation invariant $\mathcal{R}(2,5)$ code respectively, where the optimized versions of bounds (13) and (16) with respect to the parameters ρ', P, Q and the optimized version of bound (23) with respect to λ, ρ and $g(y)$ are calculated. Furthermore, the tightness of the double exponential upper bound over the new version of the generalized Shulman-Feder bound as well as the DS2 bound is clearly illustrated. In addition, even though the generalized SFB results as a special case of the DS2 bound by proper selection of the tilting measure $g(y)$ [3, Ch. 4], nevertheless for small values of the error transition probability p, the new version of the generalized version of SFB is shown to be tighter than the optimized version of the DS2 bound. This is due to the fact that in the new version of the generalized version of SFB optimization is performed with respect to all $\rho' \geq 0$, in contrary to the standard generalized SFB, where optimization is performed for $0 \leq \rho' \leq 1$.

Discussion:

1) The calculation of the coset weight distribution matrix Γ of (N,R) linear block codes is in general a computationally hard problem. Nevertheless, the full knowledge of every entry in Γ is not necessary for identifying L-list permutation invariant codes. Calculation of upper and lower bounds for the terms in the first few columns are sufficient for the previous identification. Such bounds can be deduced for example by the degree approximation function in [13, eqs. (43), (44)], the coset weight distribution inequalities by [14] or the modification of Johnson bound [15, Ch. 3], and are currently under investigation.
   L-list permutation invariant codes exist also for longer lengths. Consider for example the $[56,28,12]$ self-dual binary code of Table III in [16]. This code is 44-list permutation invariant since words of weight 8 have at most 30 codewords at Hamming distance 8, while words of Hamming weight 9 have either none or at least 45 codewords at Hamming distance 9. Furthermore, each word of weight 8 has at distance 10 more codewords than at distance 8 so that condition $\mathcal{L}_2$ is satisfied. As another example, we consider the family of extended two-error correcting BCH code of length $N=2^m$ with m even and distance matrix that of [17, Cor. 6]. Selecting $m=8$, we find that the specific code is 2367-list permutation invariant. Specifically, as is illustrated in Table II(a), all words of weight 4 have either none or at least 2368 codewords at distance 4, while all words of weight 3 have at most 45 codewords at distance 3. Moreover, condition $\mathcal{L}_2$ is satisfied since in each row of the distance matrix, there exists at least one element after column 5 strictly greater than 2368. Furthermore, let the 3-error correcting extended BCH code of length 128 with coset weight distribution matrix $\mathbf{\Gamma}$ depicted in Table II(b), [18, Table 5]. The availability of the specific distance matrix allows us to deduce that the specific code is both 1-list and 125-list permutation invariant. In more detail, every word of weight 3 has either none or 1 codeword at distance 3, while words of weight 4 have either none or at least 2 codewords of the specific code at distance 4 and at most 32 codewords at the same distance. In the meanwhile, words of weight 5 have either 0 or at least 126 at distance 5. Moreover, property $\mathcal{L}_2$ is satisfied since in each of the first two rows of the code's coset weight distribution matrix in Table II(b), an entry greater or equal than 774192 appears. This is due to the fact that the specific BCH code has 774192 codewords of weight 8, where the last parity check bit of each codeword is equal to the zero symbol.
2) Numerical simulations reveal that the double exponential upper bound as well as the SFB are rather loose for linear binary block codes that include the all ones

TABLE II
COSET WEIGHT DISTRIBUTION MATRICES $\mathbf{\Gamma}$.

(a) extended 2-error correcting BCH code of length 256

Coset Weight	Number of vectors of given weight						
	0	1	2	3	4	5	6
0	1	0	0	0	0	0	5757696
1	0	1	0	0	0	134946	0
2	0	0	1	0	2646	0	5623443
3	0	0	0	37	0	134757	0
3	0	0	0	45	0	134253	0
4	0	0	0	0	2688	0	5623296
4	0	0	0	0	2368	0	5420800
4	0	0	0	0	2880	0	5377792

(b) Extended 3-error correcting BCH code of length 128

Coset Weight	Number of vectors of given weight						
	0	1	2	3	4	5	6
0	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0
2	0	0	1	0	0	0	2667
3	0	0	0	1	0	127	0
4	0	0	0	0	2	0	2648
4	0	0	0	0	4	0	2608
4	0	0	0	0	6	0	2568
4	0	0	0	0	8	0	2528
4	0	0	0	0	0	0	2488
4	0	0	0	0	32	0	2048
5	0	0	0	0	0	126	0
6	0	0	0	0	0	0	2688

vector $\mathbf{1}$, $S_N = 1$. This is due to the fact that N is the maximizing term in the definition (20) and the previous bounds are monotonic with respect to $\mathcal{B}(\mathcal{C})$. The above restriction on the bounds' effectiveness can be circumvented by employing the Fano-Gallager bounding technique [9, eq. (16)], either on the received or the code vector space, without violating the conditions $\mathfrak{L}_1$ and $\mathfrak{L}_2$ posed for L-list permutation invariant codes.

3) Lower bounds on the weight distribution of specific codes are utilized properly in deriving upper bounds on the reliability function of binary symmetric channels for rate regions below the critical rate. Prior work towards this direction is reported among others in [19] and [20], where the exact value of the reliability function is known for specific rate regions. On the other hand, the upper bound of Theorem 2 facilitates the search for tight lower bounds on the reliability function of binary input, symmetric output discrete memoryless channels. Asymptotic upper bounds on the weight distribution of specific codes [21] are appropriate for this purpose. Specifically, the channel reliability function is defined as [2, eq. (5.8.8)]

$$E(R,p) = \lim_{N\to\infty} \frac{1}{N} \ln \frac{1}{\min_{\mathcal{C}\in\boldsymbol{C}} P_{e|\mathcal{C}}(N,R)} \tag{24}$$

where $\boldsymbol{C}$ is the set of all (N,R) codes and $P_{e|\mathcal{C}}(N,R)$ the ML error decoding probability of $\mathcal{C} \in \boldsymbol{C}$. Let a specific sequence of 1-list permutation invariant, linear binary block codes $\mathcal{C}_{\text{lb}}^N$, whose rate tends to R as N increases, while the corresponding sequence $\mathcal{B}(\mathcal{C}_{\text{lb}}^N)$, or any upper bound of it, converges

$$\lim_{N\to\infty} \mathcal{B}(\mathcal{C}_{\text{lb}}^N) = \mathcal{B}.$$

Since $\mathcal{C}_{lb}^N \subset \boldsymbol{C}$, it holds

$$\min_{\mathcal{C}\in\boldsymbol{C}} P_{e|\mathcal{C}}(N,R) \le P_{e|\mathcal{C}_{lb}^N}(N,R) \tag{25}$$

so that combining (22) with (24) and (25) we have

$$E(R,p) \ge \lim_{N\to\infty} U(R,N) \tag{26}$$

where

$$U(R,N) = \max_{\rho'\ge 0} \left(E_o(\rho') - \rho' \left(R + \mathcal{B}(\mathcal{C}_{lb}^N)\right)\right). \tag{27}$$

Then, due to the parametric analysis in [2, Sec. 5.6] and the above assumptions on the sequence of codes, $U(R,N)$ is a continuous function of $R + \mathcal{B}(\mathcal{C}_{lb}^N)$, so that from (26) we have

$$E(R,p) \ge \max_{\rho'\ge 0} \left(E_o(\rho') - \rho' \left(R + \mathcal{B}\right)\right). \tag{28}$$

Consequently, asymptotic theoretical upper bounds on the weight distribution of 1-list permutation invariant binary linear block codes constitute a useful tool in deriving tight lower bounds on the reliability function of binary input, output symmetric channels.

V. CONCLUSION

Random coding performance measures for specific codes are especially useful since they designate codes with optimum characteristics and rates close to the channel capacity. Towards this direction, the present work introduces a double exponential upper bound as well as a new version of the generalized Shulman-Feder bound on the list decoding error probability of L-list, permutation invariant, binary linear block codes, transmitted over binary input symmetric output channels. The specific analysis is motivated by the effort to smooth out the influence of the union bound effect in the calculation of the list decoding error probability of a specific code. The former bound is tighter than the generalized Shulman-Feder bound for the L-list permutation invariant codes. The latter allows the study of the performance of the previous codes with rates below the cutoff limit, and facilitates the search for tight lower bounds on the channel reliability function. Further research towards random coding bounds for specific

codes includes the study of random coding mechanisms [22] as well as the discovery of artificial ensembles of non-binary codes [6], invariant to the list decoding error probability measure, where the double exponential technique can be applied. Another interesting question is whether the current results can be extended to a more general symmetric channel setting, such as the one considered in [23, Def. 1], and closed-form expressions, analogous to those of Theorems 1 and 2, be obtained for the error decoding probability.

ACKNOWLEDGMENTS

The authors would like to thank the Associate Editor Igal Sason and the anonymous reviewer for their helpful comments that greatly enhanced the present work. Thanks are also due to the anonymous reviewer for Remark 1 and timely review.

APPENDIX A
PROOF OF LEMMA 2

Proof: Since $1/x^\rho$ is convex with respect to $x > 0$ for $\rho \geq 0$, application of Jensen's inequality to the mean value of (4) yields that

$$E\left[\Omega_L(\boldsymbol{y},\mathcal{C},\lambda,\rho)\right] \geq \frac{L^\rho \Pr(\boldsymbol{y}|\boldsymbol{c}_0)^{\lambda\rho}}{\left(E\left[\sum_{m\neq 0}\Pr(\boldsymbol{y}|\boldsymbol{c}_m,\mathcal{C})^\lambda\right]\right)^\rho} = \frac{L^\rho \Pr(\boldsymbol{y}|\boldsymbol{c}_0)^{\lambda\rho}}{\left(\sum_{m\neq 0} E\left[\Pr(\boldsymbol{y}|\boldsymbol{c}_m,\mathcal{C})^\lambda\right]\right)^\rho}. \quad (29)$$

Since the codebooks are chosen randomly from the ensemble $\mathcal{E}$, a codeword $\boldsymbol{c}_m$ of Hamming weight l in the original binary codebook $\mathcal{C}$ is permuted uniformly in $\binom{N}{l}$ distinct ways in $\mathcal{E}$ and thus

$$E\left[\Pr(\boldsymbol{y}|\boldsymbol{c}_m,\mathcal{C})^\lambda\right] = \frac{1}{\binom{N}{l}}\sum_{j=1}^{\binom{N}{l}}\Pr(\boldsymbol{y}|\boldsymbol{c}_m^j)^\lambda \quad (30)$$

where $\boldsymbol{c}_m^j$ denotes the j-th permutation of the codeword $\boldsymbol{c}_m$. Moreover, since in the original binary codebook $\mathcal{C}$, all codewords of the same Hamming weight l are permutations of each other, it holds

$$\frac{1}{\binom{N}{l}}\sum_{j=1}^{\binom{N}{l}}\Pr(\boldsymbol{y}|\boldsymbol{c}_m^j)^\lambda = \frac{1}{\binom{N}{l}}\sum_{j=1}^{\binom{N}{l}}\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^{l,j})^\lambda = E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^\lambda\right], \quad \forall m : wt_H(\boldsymbol{c}_m) = l \quad (31)$$

where $\boldsymbol{\epsilon}^l$ is an N-length binary vector of Hamming weight l. Consequently, as in [9, pp. 3048], we have

$$\left(E\left[\sum_{m\neq 0}\Pr(\boldsymbol{y}|\boldsymbol{c}_m,\mathcal{C})^\lambda\right]\right)^\rho = \left(\sum_{l=0}^N S_l E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^\lambda\right]\right)^\rho. \quad (32)$$

In order to simplify the analysis and relate the distance spectrum of the code $\mathcal{C}$ to the fully random distance spectrum, we introduce, as in [9, eq. (A12)], the probabilities v_l, b_l, $0 \leq l \leq N$, provided in (11) in Lemma 2. Then, the right hand side of (32) satisfies for $P, Q \geq 1$ and $1/P + 1/Q = 1$ and all $\rho \geq 0$

$$\left(\sum_{l=0}^N S_l E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^\lambda\right]\right)^\rho = (M-1)^\rho\left(\sum_{l=0}^N b_l\left(\frac{v_l}{b_l}\right)E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^\lambda\right]\right)^\rho \overset{(\alpha)}{\leq} (M-1)^\rho \cdot \left(\sum_{l=0}^N b_l\left(\frac{v_l}{b_l}\right)^Q\right)^{\frac{\rho}{Q}}\left(\sum_{l=0}^N b_l\left(E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^\lambda\right]\right)^P\right)^{\frac{\rho}{P}} \overset{(\beta)}{\leq} (M-1)^\rho\left(\sum_{l=0}^N b_l\left(\frac{v_l}{b_l}\right)^Q\right)^{\frac{\rho}{Q}}\left(\sum_{l=0}^N b_l E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^{\lambda P}\right]\right)^{\frac{\rho}{P}}. \quad (33)$$

In (33), (α) is due to the variant of Hölder's inequality in [2, prob. 4.15(c)], while (β) is due to Jensen's inequality for $P \geq 1$. Moreover, due to [9, eqs. (A3), (A4)], the last term of the product in the innermost term in the right hand side of (33) satisfies

$$\left(\sum_{l=0}^N b_l E_l\left[\Pr(\boldsymbol{y}|\boldsymbol{\epsilon}^l)^{\lambda P}\right]\right)^{\frac{\rho}{P}} = \left(\sum_{\boldsymbol{x}\in\mathcal{I}^N} 2^{-N}\Pr(\boldsymbol{y}|\boldsymbol{x})^{\lambda P}\right)^{\frac{\rho}{P}}. \quad (34)$$

Let

$$\lambda' = \lambda P,\ \rho' = \frac{\rho}{P},\ \lambda' = \frac{1}{1+\rho'} \quad (\lambda' \geq 0,\ \rho' \geq 0). \quad (35)$$

Then, combination of (32)-(35) yields for all $\rho' \geq 0$

$$\left(E\left[\sum_{m\neq 0}\Pr(\boldsymbol{y}|\boldsymbol{c}_m,C)^\lambda\right]\right)^\rho \leq (M-1)^{\rho' P}\cdot \left(\sum_{l=0}^N b_l\left(\frac{v_l}{b_l}\right)^Q\right)^{\frac{\rho' P}{Q}}\left(\sum_{\boldsymbol{x}\in\mathcal{I}^N} 2^{-N}\Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho'}}\right)^{\rho'}. \quad (36)$$

Replacement of (36) in (29) completes the proof. ■

APPENDIX B
PROOF OF THEOREM 1

Proof: Since the channel is memoryless, the sum in the denominator in the double exponent in (12) equals

$$\left(\sum_{\boldsymbol{x}\in\mathcal{I}^N} 2^{-N}\Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho'}}\right)^{\rho'} = 2^{-N\rho'}\left(\sum_{(x_1,\ldots,x_N)\in\mathcal{I}^N}\Pr(y_1|x_1)^{\frac{1}{1+\rho'}}\ldots\Pr(y_N|x_N)^{\frac{1}{1+\rho'}}\right)^{\rho'} = 2^{-N\rho'}\left(\sum_{x_1\in\mathcal{I}}\Pr(y_1|x_1)^{\frac{1}{1+\rho'}}\right)^{\rho'}\cdot\ldots\cdot\left(\sum_{x_N\in\mathcal{I}}\Pr(y_N|x_N)^{\frac{1}{1+\rho'}}\right)^{\rho'}. \quad (37)$$

Moreover, since the channel is symmetric, every $y_i \in \mathcal{I}$, $1 \leq i \leq N$ satisfies

$$\left(\sum_{x_i \in \mathcal{I}} \Pr(y_i|x_i)^{\frac{1}{1+\rho'}}\right)^{\rho'} = \mathcal{K}_1(p,q,\rho') \tag{38}$$

while for every $y_i \in \mathcal{J} \setminus \mathcal{I}$, $1 \leq i \leq N$, we have

$$\left(\sum_{x_i \in \mathcal{I}} \Pr(y_i|x_i)^{\frac{1}{1+\rho'}}\right)^{\rho'} = \mathcal{K}_2(p,q,\rho'). \tag{39}$$

We group together all $\boldsymbol{y}$ with $wt_H(\boldsymbol{y}) = h$ in the sum in (12), to obtain

$$\Pr(\boldsymbol{y}|\boldsymbol{c}_0) = \prod_{i=1}^{N} \Pr(y_i|0) = (1-p)^{N-h}\left(\frac{p}{q-1}\right)^h.$$

If the output alphabet is binary, $\mathcal{J} = \mathcal{I}$, then for all vectors of the above group, (37) equals

$$\left(\sum_{\boldsymbol{x} \in \mathcal{I}^N} 2^{-N}\Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho'}}\right)^{\rho'} = 2^{-N\rho'}\mathcal{K}_1(p,q,\rho')^N \tag{40}$$

and thus the upper bound in (12) satisfies Theorem 1 for $q = 2$.

Suppose $\mathcal{J} \neq \mathcal{I}$. Let i_j denote the number of times the symbol j appears in $\boldsymbol{y}$. Then, the number of vectors $\boldsymbol{y}$ with Hamming weight h, ($i_0 = N - h$) are exactly

$$\underbrace{\sum_{i_1=0}^{h} \cdots \sum_{i_{q-1}=0}^{h}}_{i_1+\ldots+i_{q-1}=h} \frac{N!}{(N-h)!i_1!\ldots i_{q-1}!}$$

since the distinct permutations of $\boldsymbol{y}$ consisting of $i_0, i_1, \ldots, i_{q-1},\ i_0 = N-h$ symbols are

$$\frac{N!}{(N-h)!i_1!\ldots i_{q-1}!}.$$

Consequently, the upper bound in (12) is equivalently expressed as

$$P_e^{\mathcal{L}} \leq \sum_{h=0}^{N}(1-p)^{N-h}\left(\frac{p}{q-1}\right)^h \underbrace{\sum_{i_1=0}^{h} \cdots \sum_{i_{q-1}=0}^{h}}_{i_1+\ldots+i_{q-1}=h} \frac{N!}{(N-h)!}\cdot$$
$$\frac{1}{i_1!\ldots i_{q-1}!}\exp\left(1-\exp\left(\frac{(M-1)^{-\rho' P}2^{N\rho'}L^{\rho' P}}{\left(\sum_{l=0}^{N} b_l\left(\frac{v_l}{b_l}\right)^Q\right)^{\frac{\rho' P}{Q}}}\cdot \right.\right.$$
$$\left.\left.\frac{\left((1-p)^{N-h}\left(\frac{p}{q-1}\right)^h\right)^{\frac{\rho'}{1+\rho'}}}{\mathcal{K}_1(p,q,\rho')^{N-h+i_1}\mathcal{K}_2(p,q,\rho')^{i_2+\ldots+i_{q-1}}} - 1\right)\right). \tag{41}$$

Rearranging the terms in the multiple sum in (41) we have

$$P_e^{\mathcal{L}} \leq \sum_{h=0}^{N}(1-p)^{N-h}\left(\frac{p}{q-1}\right)^h \sum_{i_1=0}^{h}\frac{N!}{(N-h)!}\cdot$$
$$\frac{1}{i_1!(h-i_1)!}\exp\left(1-\exp\left(\frac{(M-1)^{-\rho' P}2^{N\rho'}L^{\rho' P}}{\left(\sum_{l=0}^{N} b_l\left(\frac{v_l}{b_l}\right)^Q\right)^{\frac{\rho' P}{Q}}}\cdot \right.\right.$$
$$\left.\left.\frac{\left((1-p)^{N-h}\left(\frac{p}{q-1}\right)^h\right)^{\frac{\rho'}{1+\rho'}}}{\mathcal{K}_1(p,q,\rho')^{N-h+i_1}\mathcal{K}_2(p,q,\rho')^{h-i_1}} - 1\right)\right)\cdot$$
$$\underbrace{\sum_{i_2=0}^{h} \cdots \sum_{i_{q-1}=0}^{h}}_{i_2+\ldots+i_{q-1}=h-i_1} \frac{(h-i_1)!}{i_2!\ldots i_{q-1}!}. \tag{42}$$

Due to the multinomial expansion, the innermost term in the right hand side of (42) satisfies

$$\underbrace{\sum_{i_2=0}^{h} \cdots \sum_{i_{q-1}=0}^{h}}_{i_2+\ldots+i_{q-1}=h-i_1} \frac{(h-i_1)!}{i_2!\ldots i_{q-1}!} = (q-2)^{h-i_1}. \tag{43}$$

Combining (42) and (43), we establish Theorem 1 for $q > 2$. ■

REFERENCES

[1] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering*. Wiley, New York, 1965.

[2] R. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, New York, 1968.

[3] 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, NOW Publishers, Delft, the Netherlands, July 2006.

[4] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding," in *Proc. IEEE (ICC'93) Geneva, Switzerland*, May 1993, pp. 1064–1070.

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

[6] N. Shulman and M. Feder, "Random coding techniques for nonrandom codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2101–2104, 1999.

[7] G. D. Forney, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 206–220, March 1968.

[8] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Information and Control*, vol. 10, Part I: pp. 65–103 and Part II: pp. 522–552, Feb./May 1967.

[9] S. Shamai and I. Sason, "Variations on the Gallager bounds, connections, and applications," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3029–3051, 2002.

[10] M. Twitto, I. Sason, and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear block codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1495–1510, April 2007.

[11] C. Huffman and V. Pless, *Fundamentals of Error Correcting Codes*. Cambridge University Press, 2003.

[12] E. Hof, I. Sason, and S. Shamai, "Performance bounds for erasure, list and feedback schemes with linear block codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3754–3778, Aug. 2010.

[13] A. Cohen and N. Merhav, "Lower bounds on the error probability of block codes based on improvements on de Caen's inequality," *IEEE Trans. Inf. Theory*, vol. 50, no. 2, pp. 290–310, 2004.

[14] T. Kløve, "Bounds on the weight distribution of cosets," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2257–2260, 1996.
[15] V. Guruswami, *List Decoding of Error-Correcting Codes (Winning Thesis of the 2002 ACM Doctoral Dissertation Competition)*, ser. Lecture Notes in Computer Science. Springer, 2004, vol. 3282.
[16] M. Ozeki, "On covering radii and coset weight distributions of extremal binary self–dual codes of length 56," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2359–2372, Nov. 2000.
[17] P. Charpin, "Weight distributions of cosets of two-error-correcting binary BCH codes, extended or not," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1425–1442, Sep. 1994.
[18] P. Charpin and V. A. Zinoviev, "On coset weight distributions of the 3-error-correcting BCH codes," *SIAM J. Discrete Math.*, vol. 10, no. 1, pp. 128–145, Feb. 1997.
[19] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 385–398, 1999.
[20] Y. Ben-Haim and S. Litsyn, "Improved upper bounds on the reliability function of the Gaussian channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 5–12, 2008.
[21] A. Ashikhmin, A. Barg, and S. Litsyn, "Estimates of the distance distribution of codes and designs," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1050–1061, 2001.
[22] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete-memoryless channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 417–438, March 2004.
[23] E. Hof, I. Sason, and S. Shamai, "Performance bounds for non-binary linear block codes over memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 977–996, March 2009.

Kostis Xenoulis (S'08) received the B.Sc. degree in informatics & telecommunications, the M.Sc. degree in signal processing for telecommunications and multimedia and the PhD degree from the University of Athens, Greece, in 2003, 2005 and 2010 respectively. His research interests are in the area of information theory.

Nicholas Kalouptsidis (M'82-SM'85) received the B.Sc. degree in mathematics from the University of Athens in 1973 and the Ph.D. degree in systems science and mathematics from Washington University, St. Louis, MO, in 1976.

He has held visiting positions with Washington University; Politecnico di Torino; Northeastern University, Boston, MA; CNET, Lannion, France; and University of Utah, Salt Lake City. In spring 2008, he was a visiting scholar with Harvard University. He is currently a Professor with the Department of Informatics and Telecommunications, University of Athens. He is the author of the textbook Signal Processing Systems: Theory and Design (New York: Wiley, 1997) and coeditor, with S. Theodoridis, of the book Adaptive System Identification and Signal Processing Algorithms (Englewood Cliffs, NJ: Prentice-Hall, 1993). His research interests are in system theory and signal processing.