An improved bound on the list error probability and list distance properties

Bocharova, Irina; Kudryashov, Boris; Johannesson, Rolf; Loncar, Maja

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/TIT.2007.911176

2008

An Improved Bound on the List Error Probability and List Distance Properties

Irina E. Bocharova, Rolf Johannesson, *Fellow, IEEE*,
Boris D. Kudryashov, and Maja Lončar, *Student Member, IEEE*

***Abstract*— List decoding of binary block codes for the additive white Gaussian noise channel is considered. The output of a list decoder is a list of the L most likely codewords, that is, the L signal points closest to the received signal in the Euclidean-metric sense. A decoding error occurs when the transmitted codeword is not on this list. It is shown that the list error probability is fully described by the so-called list configuration matrix, which is the Gram matrix obtained from the signal vectors forming the list. The worst-case list configuration matrix determines the minimum list distance of the code, which is a generalization of the minimum distance to the case of list decoding. Some properties of the list configuration matrix are studied and their connections to the list distance are established. These results are further exploited to obtain a new upper bound on the list error probability, which is tighter than the previously known bounds. This bound is derived by combining the techniques for obtaining the tangential union bound with an improved bound on the error probability for a given list. The results are illustrated by examples.**

***Index Terms*— List configuration matrix, list decoding, list distance, list error probability, tangential union bound**

I. Introduction

The optimal decoding method that minimizes the sequence error probability at the receiver is maximum *a posteriori* probability (MAP) sequence decoding, which reduces to maximum-likelihood (ML) decoding when all the sequences (codewords) are *a priori* equiprobable. When signalling over the additive white Gaussian noise (AWGN) channel, ML decoding is equivalent to finding the codeword with the smallest Euclidean distance from the received sequence.

List decoding, introduced in [1], is a generalization of ML decoding—a list decoder is not restricted to find a single estimate of the transmitted codeword but delivers a list of most likely codewords, closest to the received word in terms of a given metric. Decoding is considered successful if the transmitted codeword is included in the list.

List decoding has found applications in concatenated coding schemes, often used in combination with automatic-repeat-request (ARQ) strategies: The outer error detection code, such as the cyclic redundancy check (CRC) code, is combined with an inner error correcting code. At the receiver, rather than using the ML decoder to decode the inner code, a list decoder may be employed to find a list of the most probable sequences, which are subsequently checked by the outer CRC decoder. Only if none of the sequences on the list satisfies the CRC parity constraints, a retransmission is requested via a feedback channel. This scenario was investigated in [2], where the list-Viterbi algorithm was developed for list decoding of convolutional codes. It was shown in [2] that already moderate list sizes provide significantly lower error probability than decoding with list size equal to one. Similar applications of list decoding for speech recognition (where the outer CRC code is ”replaced” by a language processor) were investigated in [3] where the search for the list of sequences is performed with the tree-trellis algorithm (TTA), *cf.* also [4].

Since the introduction of turbo codes [5] more than a decade ago, iterative decoding and detection algorithms have received much attention. Iterative (turbo) schemes bypass the prohibitively complex optimal decoding of the overall concatenated code by employing simpler constituent decoders as separate entities which iteratively exchange soft information on decoded symbols. Constituent soft-input soft-output (SISO) decoders can be realized with the BCJR algorithm [6]; however, its complexity becomes prohibitively high for codes with large trellis state space. This is typically the case when the constituent codes are block codes, as is the case in the product codes. In this context, *list-based* SISO decoders have been recently proposed as a low-complexity alternative to the BCJR decoding (*cf.* [7], [8], and the references therein). These decoding methods use a list of candidate codewords and their metrics to compute approximate symbol reliabilities. In [8], the list is obtained by the bidirectional efficient algorithm for searching code trees (BEAST) and it was demonstrated that a list of only a few most probable codewords suffices for accurate estimation of symbol reliabilities. More generally, the turbo receiver principle is applicable to many communication systems that can be represented as concatenated coding schemes, where the inner code is, for example, realized by a modulator followed by the intersymbol interference (ISI) channel, or by a space-time code for multiple-antenna transmission. List-based inner SISO detectors that form the symbol reliability estimates from a list of signal vectors were proposed, *e.g.*, in [9] for MIMO transmission and in [10] for ISI equalization.

Although different applications of list decoding have been considered in a large number of papers, only a few papers were devoted to estimating the list error probability. Since

This research was supported in part by the Royal Swedish Academy of Sciences in cooperation with the Russian Academy of Sciences and in part by the Swedish Research Council under Grant 621-2004-4703. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Seattle, USA, July 2006.

I. E. Bocharova and B. D. Kudryashov are with the Department of Information Systems, St. Petersburg University of Aerospace Instrumentation, St. Petersburg, 190000, Russia (e-mail: irina@it.lth.se; boris@it.lth.se).

R. Johannesson and M. Lončar are with the Department of Electrical and Information Technology, Lund University, P.O. Box 118, SE-22100 Lund, Sweden (e-mail: rolf@eit.lth.se; maja@eit.lth.se).

exact expressions for the error probability are most often not analytically tractable, tight bounds are useful tools for estimating the performance of the system and for identifying the parameters that dominate its behavior. The earliest results regarding list decoding were obtained for code *ensembles*, using random coding arguments: Bounds on the asymptotic rates of binary block codes used with list decoding over the binary symmetric channel were investigated in [11], [12], and later also in [13] (for more references to related works see [13]). The asymptotic behavior of the list error probability was analyzed in [14] and [15], where bounds on the error exponents were obtained. More recently, asymptotic bounds on the code size and the error exponents for list decoding in Euclidean space were derived in [16].

When estimating the error rate performance of ML decoding (that is, list decoding with the list of size one) for a *specific* code used for communicating over the AWGN channel, the most commonly used upper bound is obtained by applying the Bonferroni-type inequality which states that the probability of a union of events is not larger than the sum of the probabilities of the individual events[1]. This yields the well-known union bound, which upper-bounds the error probability by the sum of pairwise error events. This bound is simple to compute and requires only the knowledge of the code spectrum; however, it is tight only at high signal-to-noise ratio (SNR) levels, while at low and moderate SNRs it becomes too loose due to the fact that in the sum of pairwise error probabilities, the same error event may be counted many times. There have been several improvements during the past two decades that yield much tighter bounds than the union bound. These include the tangential bound (TB) [17], the sphere bound (SB) [18], [19], and the tangential-sphere bound (TSB) [20]. These bounds are based on the well-known bounding principle introduced by Fano [21] for random codes and adapted by Gallager [22] for specific codes, where the received signal space is partitioned into two disjoint regions, $\mathcal{R}$ and its complement $\mathcal{R}^{\mathrm{c}}$, of few and many errors, respectively. The error probability $\Pr(\varepsilon)$ is thus split into the sum of two error probabilities, when the received signal $\boldsymbol{r}$ resides inside and outside $\mathcal{R}$, that is, $\Pr(\varepsilon) = \Pr(\varepsilon, \boldsymbol{r} \in \mathcal{R}) + \Pr(\varepsilon, \boldsymbol{r} \in \mathcal{R}^{\mathrm{c}})$. The first term, referring to the region of few errors can be upper-bounded using a *union bound*, while the second term is bounded simply by $\Pr(\boldsymbol{r} \in \mathcal{R}^{\mathrm{c}})$. We call this principle the Fano-Gallager bounding; it is also referred to as the Gallager-Fano bound [23] or Gallager's first bounding principle [24]. The TB for equi-energy signals (which all lie on a hypersphere) was derived by splitting the noise vector into radial and tangential components, which lie along and perpendicular to the transmitted signal, respectively. The few-error region $\mathcal{R}$ in the TB is a half-space where the magnitude of radial noise is not larger than a certain threshold. The SB is obtained by considering the spherical region $\mathcal{R}$, and finally, in the TSB, which is tighter than the previous bounds, both approaches are combined and the region $\mathcal{R}$ is a circular cone with the axis passing through the transmitted signal point. A detailed treatment and comparisons of various Fano-Gallager-type bounds can be found in [23] and [24]. Recently, two new bounds that improve upon the TSB have been proposed: the so-called added-hyperplane (AHP) bound [25] and the improved TSB (ITSB) [26]. Both bounds are obtained by upper-bounding the probability $\Pr(\varepsilon, \boldsymbol{r} \in \mathcal{R})$ using a tighter, *second-order* Bonferroni-type inequality[2] instead of the union bound used in the TSB.

Generalization of the bounds for ML decoding to list decoding is not straightforward. A list error event is defined with respect to a list of L codewords, which implies that the pairwise error events considered in ML decoding translate to $(L+1)$-wise list-error events. Geometrical properties of list configurations were investigated in [27], and used to derive a union bound on the list error probability. The notions of the Euclidean and Hamming list distances were introduced and it was shown that these distances are generalizations of the Euclidean and Hamming distances of the code.

In this paper, we build upon the work of [27] and investigate the properties of the so-called list distance and its relations to the list configurations. Moreover, using the tangential-bound approach from [17], we improve the union bound of [27]. Similarly as in [28], we first derive a tighter bound on the error probability for a *given list*, and then obtain a new upper bound on the list error probability by combining this tighter bound with a modified tangential bound.

II. Geometrical Aspects of List Decoding

In this section the notions of the *list distance* and the *list configuration matrix* are introduced and their properties and relations are established. The results presented in the first two subsections have mostly appeared in [27]; however, we present them here in an extended form, supported by examples and more detailed discussion. In the last subsection, the relation of the list distance and the *average radius* introduced in [12] is discussed. This section serves as a basis for the derivation of the list-error probability bounds presented in Sections III and IV.

A. List Decoding

Let $\mathcal{S}=\{\boldsymbol{s}_i\}$, $i=0,1,...,M-1$, be an arbitrary constellation of $|\mathcal{S}|=M$ equiprobable signal points $\boldsymbol{s}_i=(s_i^{(1)}s_i^{(2)}\ldots s_i^{(N)})$ used to communicate over an additive white Gaussian noise

[1]The probability of a union of events E_i, $i = 1,2,\ldots,M$ is equal to $\Pr\left(\bigcup_{i=1}^{M} E_i\right) = \sum_{i=1}^{M}\Pr(E_i) - \sum_{1\le i_1<i_2\le M}\Pr(E_{i_1}\cap E_{i_2}) + ... + (-1)^{M+1}\sum_{1\le i_1<i_2<\ldots<i_M\le M}\Pr(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_M})$. Truncating the right-hand side expression after the first term yields an upper bound referred to as the Bonferroni inequality of the *first order*, since it depends only on the probabilities of elementary events.

[2]The probability of a union of events $E_i, i = 1,2,\ldots,M$ can be expressed as $\Pr\left(\bigcup_{i=1}^{M} E_i\right) = \Pr(E_1) + \Pr(E_2\cap E_1^{\mathrm{c}}) + \Pr(E_3\cap E_1^{\mathrm{c}}\cap E_2^{\mathrm{c}}) + ... + \Pr(E_M\cap E_1^{\mathrm{c}}\cap E_2^{\mathrm{c}}\cap...\cap E_{M-1}^{\mathrm{c}})$, where E_i^{c} denotes the complement of E_i. From here follows an upper bound used in [25]: $\Pr\left(\bigcup_{i=1}^{M} E_i\right) \le \Pr(E_1) + \Pr(E_2\cap E_1^{\mathrm{c}}) + \Pr(E_3\cap E_{j_1}^{\mathrm{c}}) + ... + \Pr(E_M\cap E_{j_{M-2}}^{\mathrm{c}})$, where the tightness of the bound is determined by the ordering of the events and the choices of indices $j_1 \in \{1,2\}$, $j_2 \in \{1,2,3\}$, $\ldots$, $j_{M-2} \in \{1,2,...,M-1\}$. This bound is a Bonferroni inequality of the *second order*, since it involves pairwise joint event probabilities.

(AWGN) channel. Assume that the N-tuple $\boldsymbol{s}_0$ was transmitted. The discrete-time received signal is

$$\boldsymbol{r} = \boldsymbol{s}_0 + \boldsymbol{n} \tag{1}$$

where the noise vector $\boldsymbol{n}$ consists of independent zero-mean Gaussian random variables with variance $N_0/2$. We say that an error for a given list $\mathcal{L} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_L\}$, of size $L < M$, occurs if $\boldsymbol{s}_0 \notin \mathcal{L}$, which implies that $\boldsymbol{s}_0$ is further from the received signal than all signals on the list, that is,

$$d_{\mathrm{E}}^2(\boldsymbol{r}, \boldsymbol{s}_\ell) \le d_{\mathrm{E}}^2(\boldsymbol{r}, \boldsymbol{s}_0), \quad \ell = 1, 2, \ldots, L \tag{2}$$

where $d_{\mathrm{E}}^2(\boldsymbol{r}, \boldsymbol{s}_\ell) = ||\boldsymbol{r} - \boldsymbol{s}_\ell||^2$ is the squared Euclidean distance between the received vector $\boldsymbol{r}$ and the vector $\boldsymbol{s}_\ell$ from the list. This is slightly pessimistic, since it implies that when (2) is fulfilled with equality, we always include the erroneous N-tuple in the list. By substituting (1) into (2) we obtain

$$\|\boldsymbol{n} + \boldsymbol{s}_0 - \boldsymbol{s}_\ell\|^2 \le \|\boldsymbol{n}\|^2, \quad \ell = 1, 2, ..., L$$

which is equivalent to

$$\langle \boldsymbol{n}, \boldsymbol{s}_\ell - \boldsymbol{s}_0 \rangle \ge d_{\mathrm{E}}^2(\boldsymbol{s}_0, \boldsymbol{s}_\ell)/2, \quad \ell = 1, 2, \ldots, L \tag{3}$$

where $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$ denotes the inner product of the row vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

Now let t_ℓ denote the inner product $t_\ell = \langle \boldsymbol{n}, \boldsymbol{s}_\ell - \boldsymbol{s}_0 \rangle$, $\ell = 1, 2, \ldots, L$. Then the vector $\boldsymbol{t} = (t_1\, t_2\, \ldots\, t_L)$ is a Gaussian random vector with zero mean and covariance matrix

$$\mathrm{E}\left[\boldsymbol{t}^{\mathrm{T}}\boldsymbol{t}\right] = \frac{N_0}{2}\boldsymbol{\Gamma}.$$

The entries of the $L \times L$ matrix $\boldsymbol{\Gamma} = \{\gamma_{ij}\}$, $i, j = 1, 2, \ldots, L$, are

$$\gamma_{ij} = \langle \boldsymbol{s}_i - \boldsymbol{s}_0, \boldsymbol{s}_j - \boldsymbol{s}_0 \rangle = (d_{\mathrm{E}0i}^2 + d_{\mathrm{E}0j}^2 - d_{\mathrm{E}ij}^2)/2 \tag{4}$$

where $d_{\mathrm{E}ij} = d_{\mathrm{E}}(\boldsymbol{s}_i, \boldsymbol{s}_j) = ||\boldsymbol{s}_i - \boldsymbol{s}_j||$. Thus, $\boldsymbol{\Gamma}$ is the Gram matrix of the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, $\ell = 1, 2, \ldots, L$. We call $\boldsymbol{\Gamma}$ the *list configuration matrix*.

Let $\boldsymbol{\gamma}$ denote the vector of the main-diagonal elements of the list configuration matrix $\boldsymbol{\Gamma}$, that is,

$$\boldsymbol{\gamma} = \left(d_{\mathrm{E}01}^2\ d_{\mathrm{E}02}^2\ \ldots\ d_{\mathrm{E}0L}^2\right).$$

From (3) it follows that the list error probability for any list with given configuration matrix $\boldsymbol{\Gamma}$ is given by[3]

$$P_{\mathrm{e}L}(\boldsymbol{\Gamma}) = \Pr(\boldsymbol{t} \ge \boldsymbol{\gamma}/2). \tag{5}$$

Consider a binary $(N, K, d_{\mathrm{H}\min})$ block code $\mathcal{C} = \{\boldsymbol{v}_i\}$, $i = 0, 1, \ldots, 2^K - 1$, of length N, dimension K, and minimum distance $d_{\mathrm{H}\min}$. Since the distance spectrum is a property of a linear code, we will hereinafter assume code linearity, although this condition is not necessary for the results presented in this section. When the code $\mathcal{C}$ is used with binary phase shift keying (BPSK) to communicate over an AWGN channel, the binary code symbols $v_i^{(j)} \in \{0, 1\}$, $j = 1, 2, \ldots, N$, are mapped onto the symbols

$$s_i^{(j)} = (1 - 2v_i^{(j)})\sqrt{E_s}$$

[3]Hereinafter, the relation $\boldsymbol{a} \ge \boldsymbol{b}$ between two vectors of the same length L should be interpreted element-wise, that is, $\boldsymbol{a} \ge \boldsymbol{b}$ holds if and only if $a_i \ge b_i$, $\forall\, i = 1, 2, \ldots, L$.

where E_s is the symbol energy equal to $E_s = E_b R$, where $R = K/N$ is the code rate and E_b is the energy per bit. All the signal points $\boldsymbol{s}_i$, $i = 0, 1, ..., 2^K - 1$, have the same energy, $||\boldsymbol{s}_i|| = \sqrt{NE_s}$, that is, they lie on a hypershere of radius $\sqrt{NE_s}$ in the N-dimensional Euclidean space. The squared Euclidean distance between two signal points is proportional to the Hamming distance between the corresponding codewords, that is,

$$d_{\mathrm{E}}^2(\boldsymbol{s}_i, \boldsymbol{s}_j) = 4E_s d_{\mathrm{H}}(\boldsymbol{v}_i, \boldsymbol{v}_j). \tag{6}$$

The minimum squared Euclidean distance of the code is

$$d_{\mathrm{E}\min}^2 = \min_{\boldsymbol{s}_i \ne \boldsymbol{s}_j} \{d_{\mathrm{E}}^2(\boldsymbol{s}_i, \boldsymbol{s}_j)\} = 4E_s d_{\mathrm{H}\min}. \tag{7}$$

Then the list configuration matrix can be written as $\boldsymbol{\Gamma} = 4E_s\boldsymbol{\Gamma}_{\mathrm{H}}$, where $\boldsymbol{\Gamma}_{\mathrm{H}}$ is the *normalized list configuration matrix* whose entries are

$$\gamma_{\mathrm{H}ij} = (d_{\mathrm{H}0i} + d_{\mathrm{H}0j} - d_{\mathrm{H}ij})/2 \tag{8}$$

where $d_{\mathrm{H}ij} = d_{\mathrm{H}}(\boldsymbol{v}_i, \boldsymbol{v}_j)$. Examples of $\boldsymbol{\Gamma}_{\mathrm{H}}$ for the $(7, 4, 3)$ Hamming code, the $(8, 4, 4)$ extended Hamming code, and the $(24, 12, 8)$ extended Golay code are given in Tables I, II, and III, respectively. Without loss of generality, we assume that the reference signal $\boldsymbol{s}_0$ corresponds to the all-zero codeword $\boldsymbol{v}_0 = \boldsymbol{0}$. Each row in Tables I–III corresponds to a distinct value of the Hamming list distance $d_{\mathrm{H}L}$ which will be explained in the next subsection. Several configuration matrices can yield the same $d_{\mathrm{H}L}$. For each list size L, list configurations are listed in the order of increasing $d_{\mathrm{H}L}$. The last column in the tables shows the number of lists $N(\boldsymbol{\Gamma})$ with the same list configuration matrix $\boldsymbol{\Gamma}$. Note that the ordering of the codewords on the list is irrelevant. Hence, for a given list configuration, $N(\boldsymbol{\Gamma})$ is the number of combinations with matrices $\boldsymbol{\Gamma}$ that are equal up to a permutation of the main-diagonal and the corresponding off-diagonal entries.

For list size $L = 1$, the normalized list configuration matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ reduces to codeword's Hamming weight $d_{\mathrm{H}0i}$ and values of $N(\boldsymbol{\Gamma})$ yield the distance spectrum of a code.

For list size $L = 2$, consider, for example, the $(7, 4, 3)$ Hamming code from Table I. There are six possible weight combinations to form a list of two codewords, corresponding to six list configuration matrices $\boldsymbol{\Gamma}_{\mathrm{H}}$. Consider, for example, lists with two minimum-weight codewords, $d_{\mathrm{H}\min} = 3$. There are $N(\boldsymbol{\Gamma}) = \binom{7}{2} = 21$ such lists. All the minimum-weight codewords of the Hamming code have the pairwise distance $d_{\mathrm{H}ij} = 4$. Hence, the corresponding normalized list configuration matrix is $\boldsymbol{\Gamma}_{\mathrm{H}} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Next, consider lists-of-two that contain one codeword of weight 3 and one of weight 4. There are in total $7 \times 7 = 49$ such pairs, out of which $N(\boldsymbol{\Gamma}) = 42$ pairs have pairwise distance 3 and hence their configuration matrix is $\boldsymbol{\Gamma}_{\mathrm{H}} = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$. The remaining 7 pairs are at the distance 7 and their configuration matrix is $\boldsymbol{\Gamma}_{\mathrm{H}} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$.

In [27] the following union-type bound on the list error probability for a given list size L was derived

$$P_{\mathrm{e}L} \le \sum_{\boldsymbol{\Gamma}} N(\boldsymbol{\Gamma}) P_{\mathrm{e}L}(\boldsymbol{\Gamma}) \tag{9}$$

TABLE I

LIST CONFIGURATIONS FOR THE $(7, 4, 3)$ HAMMING CODE

L	$\mathbf{\Gamma}_{\mathrm{H}} = \mathbf{\Gamma}/(4E_s)$	$d_{\mathrm{H}L}$	$N(\mathbf{\Gamma})$
1	(3)	3	7
	(4)	4	7
	(7)	7	1
2	$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$	4.50	21, 42
	$\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$	5.33	21
	$\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 3 \\ 3 & 7 \end{pmatrix}, \begin{pmatrix} 4 & 4 \\ 4 & 7 \end{pmatrix}$	7	7, 7, 7
3	$\begin{pmatrix} 3 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 4 \end{pmatrix}$	5	105
	$\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$	5.40	35, 105
	$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$	6	35
	$\begin{pmatrix} 3 & 1 & 3 \\ 1 & 3 & 3 \\ 3 & 3 & 7 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 7 \end{pmatrix}, \begin{pmatrix} 4 & 2 & 4 \\ 2 & 4 & 4 \\ 4 & 4 & 7 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 2 & 4 \end{pmatrix}$	7	21, 42, 42, 21, 42
	$\begin{pmatrix} 3 & 0 & 3 \\ 0 & 4 & 4 \\ 3 & 4 & 7 \end{pmatrix}$	∞	7
4	$\begin{pmatrix} 3 & 1 & 1 & 2 \\ 1 & 3 & 1 & 2 \\ 1 & 1 & 3 & 2 \\ 2 & 2 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 2 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 2 & 4 & 2 \\ 2 & 2 & 2 & 4 \end{pmatrix}$	5.50	140, 210
	$\begin{pmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ 2 & 2 & 4 & 2 \\ 2 & 2 & 2 & 4 \end{pmatrix}$	6	35, 140
	$\vdots$	$\vdots$	$\vdots$

TABLE II

LIST CONFIGURATIONS FOR THE $(8, 4, 4)$ EXTENDED HAMMING CODE

L	$\mathbf{\Gamma}_{\mathrm{H}} = \mathbf{\Gamma}/(4E_s)$	$d_{\mathrm{H}L}$	$N(\mathbf{\Gamma})$
1	(4)	4	14
	(8)	8	1
2	$\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$	5.33	84
	$\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 4 \\ 4 & 8 \end{pmatrix}$	8	7, 14
3	$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$	6	280
	$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 2 & 4 \\ 2 & 4 & 4 \\ 4 & 4 & 8 \end{pmatrix}$	8	84, 84
	$\begin{pmatrix} 4 & 0 & 4 \\ 0 & 4 & 4 \\ 4 & 4 & 8 \end{pmatrix}$	∞	7

where $N(\mathbf{\Gamma})$ is the number of lists of size L which have the same list configuration matrix $\mathbf{\Gamma}$. It follows from (5) and (9) that the list error probability can be fully described in terms of the properties of the Gram matrix $\mathbf{\Gamma}$. For binary codes, this matrix determines the so-called minimum Hamming list distance of the code [27], $d_{\mathrm{H}L\min}$, which plays the same role for list decoding as the minimum distance for maximum-likelihood decoding. In the next subsection, the list distance is defined and illustrated by examples.

B. List Radius and List Distance

Consider first maximum-likelihood decoding, that is, list decoding with list size $L = 1$. The largest contribution to the error probability is obtained when the received point $\boldsymbol{r}$ is exactly between the two closest signal points, that is, signal points at the minimum Euclidean distance $d_{\mathrm{E}\min}$. Thus, $\boldsymbol{r}$ is the center of this constellation of $L + 1 = 2$ signal points. Next, we generalize this approach to arbitrary list size L.

TABLE III
LIST CONFIGURATIONS FOR THE $(24, 12, 8)$ EXTENDED GOLAY CODE

L	$\mathbf{\Gamma}_{\rm H} = \mathbf{\Gamma}/(4E_s)$	$d_{{\rm H}L}$	$N(\mathbf{\Gamma})$
1	(8)	8	759
	(12)	12	2576
	(16)	16	759
	(24)	24	1
2	$\begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$	10.67	106260
	$\begin{pmatrix} 8 & 2 \\ 2 & 8 \end{pmatrix}, \begin{pmatrix} 8 & 6 \\ 6 & 12 \end{pmatrix}$	12.80	170016, 340032
	$\begin{pmatrix} 8 & 4 \\ 4 & 12 \end{pmatrix}, \begin{pmatrix} 12 & 8 \\ 8 & 12 \end{pmatrix}$	14.40	1275120, 63760
	$\begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}, \begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix}, \begin{pmatrix} 8 & 8 \\ 8 & 16 \end{pmatrix}$	16	11385, 2040192, 22770
	$\vdots$	$\vdots$	$\vdots$

Let $\mathcal{S}$ be an arbitrary constellation of $|\mathcal{S}| = M$ signal points in the Euclidean space, and let $\mathcal{S}_L$ be an arbitrary subset $\mathcal{S}_L = \{\boldsymbol{s}_0, \boldsymbol{s}_1, ..., \boldsymbol{s}_L\} \subseteq \mathcal{S}$ of $L+1$ signal points. Then, the *minimum list radius* of the constellation $\mathcal{S}$, for a list size L, is defined as

$$R_{L\,\min} \triangleq \min_{\mathcal{S}_L} \min_{\boldsymbol{r}} \max_{0 \le k \le L} \{d_{\rm E}(\boldsymbol{s}_k, \boldsymbol{r})\}. \tag{10}$$

For a given signal subset $\mathcal{S}_L$, the list radius is the radius of the smallest sphere $\mathscr{S}$ that contains (encompasses) the points of $\mathcal{S}_L$ (that is, the points lie *on* or *inside* the sphere), and the minimizing $\boldsymbol{r}$ is the center of this sphere. Minimization over all possible subsets $\mathcal{S}_L \subseteq \mathcal{S}$ yields the smallest list radius for a given list size L. Thus, if the noise $\boldsymbol{n}$ is such that the received signal point $\boldsymbol{r}$ is closer than $R_{L\,\min}$ to the transmitted signal point, it is guaranteed that the transmitted signal point will be among the L points closest to the received signal and the list decoder will not make an error. Clearly, the list radius is the distance from the transmitted signal point to the closest point of the list error decision region. Like the minimum distance, the minimum list radius is also a constellation property.

Let $\boldsymbol{s}_0 \in \mathcal{S}_L$ be the reference (transmitted) signal, and let $\mathbf{\Gamma}$ be the list configuration matrix corresponding to the signal subset $\mathcal{S}_L$. Assume that the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, $\ell = 1, 2, ..., L$ are linearly independent; then the matrix $\mathbf{\Gamma}$ has full rank. The following theorem from [27] specifies the center and the radius of the *circumsphere* $\widetilde{\mathscr{S}}$ of the set $\mathcal{S}_L$ (that is, the sphere such that all the points of $\mathcal{S}_L$ lie on the sphere). We also present the proof, in an extended form, as some of its steps will prove useful later on.

Theorem 1: Let $\mathcal{S}_L = \{\boldsymbol{s}_0, \boldsymbol{s}_1, ..., \boldsymbol{s}_L\}$ be a set of $L+1$ signal points such that the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, $\ell = 1, 2, \ldots, L$, are linearly independent. Let $\mathbf{\Gamma}$ be the corresponding Gram matrix of the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, and let $\boldsymbol{\gamma}$ be the row vector of its main-diagonal elements. Then the radius $\widetilde{R}_L$ of the circumsphere $\widetilde{\mathscr{S}}$ of $\mathcal{S}_L$ is given by

$$\widetilde{R}_L(\mathbf{\Gamma}) = \frac{1}{2}\sqrt{\boldsymbol{\gamma}\mathbf{\Gamma}^{-1}\boldsymbol{\gamma}^{\rm T}} \tag{11}$$

and the center $\boldsymbol{\rho}$ of the sphere $\widetilde{\mathscr{S}}$ is given by

$$\boldsymbol{\rho} - \boldsymbol{s}_0 = \frac{1}{2}\boldsymbol{\gamma}\mathbf{\Gamma}^{-1}\boldsymbol{S}, \qquad \boldsymbol{S} = \begin{pmatrix} \boldsymbol{s}_1 - \boldsymbol{s}_0 \\ \boldsymbol{s}_2 - \boldsymbol{s}_0 \\ \vdots \\ \boldsymbol{s}_L - \boldsymbol{s}_0 \end{pmatrix}. \tag{12}$$

Proof: Since all the points of $\mathcal{S}_L$ lie on the sphere $\widetilde{\mathscr{S}}$, its radius $\widetilde{R}_L$ satisfies

$$\widetilde{R}_L(\mathbf{\Gamma}) = ||\boldsymbol{\rho} - \boldsymbol{s}_0|| = ||\boldsymbol{\rho} - \boldsymbol{s}_\ell||, \quad \ell = 1, 2, \ldots, L. \tag{13}$$

From here it follows that

$$2\,\langle \boldsymbol{\rho} - \boldsymbol{s}_0, \boldsymbol{s}_\ell - \boldsymbol{s}_0 \rangle = ||\boldsymbol{s}_\ell - \boldsymbol{s}_0||^2 \tag{14}$$

which can be rewritten in vector form as

$$2(\boldsymbol{\rho} - \boldsymbol{s}_0)\boldsymbol{S}^{\rm T} = \boldsymbol{\gamma} \tag{15}$$

where $\boldsymbol{S}$ is given by (12). Note that $\boldsymbol{S}\boldsymbol{S}^{\rm T} = \mathbf{\Gamma}$.

Now let $\boldsymbol{\zeta} = (\zeta_1\, \zeta_2\, \ldots\, \zeta_L)$ be a vector of coefficients of the decomposition of the vector $\boldsymbol{\rho} - \boldsymbol{s}_0$ in the L-dimensional basis consisting of the linearly independent vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$. Then we can write

$$\boldsymbol{\rho} - \boldsymbol{s}_0 = \zeta_1(\boldsymbol{s}_1 - \boldsymbol{s}_0) + \zeta_2(\boldsymbol{s}_2 - \boldsymbol{s}_0) + \cdots + \zeta_L(\boldsymbol{s}_L - \boldsymbol{s}_0)$$

or, equivalently,

$$\boldsymbol{\rho} - \boldsymbol{s}_0 = \boldsymbol{\zeta}\boldsymbol{S}. \tag{16}$$

Substituting (16) into (13) yields

$$\widetilde{R}_L(\mathbf{\Gamma}) = \sqrt{\|\boldsymbol{\rho} - \boldsymbol{s}_0\|^2} = \sqrt{\boldsymbol{\zeta}\boldsymbol{S}\boldsymbol{S}^{\rm T}\boldsymbol{\zeta}^{\rm T}} = \sqrt{\boldsymbol{\zeta}\mathbf{\Gamma}\boldsymbol{\zeta}^{\rm T}}. \tag{17}$$

From (15) and (16) it follows that $2\boldsymbol{\zeta}\mathbf{\Gamma} = \boldsymbol{\gamma}$, which yields

$$\boldsymbol{\zeta} = \frac{1}{2}\boldsymbol{\gamma}\mathbf{\Gamma}^{-1}. \tag{18}$$

Substituting (18) into (16) and (17) yields (12) and (11), respectively, which completes the proof.

■

Clearly, if the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$ are linearly independent, the sphere $\widetilde{\mathscr{S}}$ is L-dimensional. If, however, some of the vectors are linearly dependent, the Gram matrix $\mathbf{\Gamma}$ is singular, that is, $\det(\mathbf{\Gamma}) = 0$, and the radius is $\widetilde{R}_L = \infty$, since the signal points $\boldsymbol{s}_\ell$ lie in a reduced subspace (*cf.* Examples 3–4 below).

For a given signal set $\mathcal{S}_L$, the circumsphere $\widetilde{\mathscr{S}}$, may, in general, not be the smallest sphere that encompasses the points of $\mathcal{S}_L$ (which is the sphere that determines the list radius).

Let $\mathscr{S}$ denote the smallest encompassing sphere of the set $\mathcal{S}_L$, such that the reference point $\boldsymbol{s}_0$ lies on the sphere, and the remaining points $\boldsymbol{s}_\ell$, $\ell = 1, 2, ..., L$, lie either on or inside $\mathscr{S}$ (more precisely, at least one more point $\boldsymbol{s}_\ell \in \mathcal{S}_L$, other than $\boldsymbol{s}_0$, will lie on such a sphere). Assume that the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$ are linearly independent. Then it was shown in [27] that the radius R_L of the sphere $\mathscr{S}$ is given by

$$R_L(\boldsymbol{\Gamma}) = \max_{\mathcal{I}\,:\,\boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1}\geq\mathbf{0}} \left\{ \frac{1}{2}\sqrt{\boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1}\boldsymbol{\gamma}_\mathcal{I}^{\mathrm{T}}} \right\} \tag{19}$$

where the maximization is performed over all signal subsets $\mathcal{I} \subseteq \mathcal{S}_L$ that contain the reference point $\boldsymbol{s}_0$, such that their corresponding configuration matrix $\boldsymbol{\Gamma}_\mathcal{I}$ and its main-diagonal vector $\boldsymbol{\gamma}_\mathcal{I}$ fulfill the condition $\boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1} \geq \mathbf{0}$. Note that $\boldsymbol{\Gamma}_\mathcal{I}$ is a main submatrix of the configuration matrix $\boldsymbol{\Gamma}$, obtained by deleting those rows and columns that correspond to the signal points not included in the chosen subset $\mathcal{I}$.

Theorem 1 and formula (19) imply the following:

- The list radius R_L is the largest radius of the circumspheres of all the signal subsets $\mathcal{I}$ such that $\boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1} \geq \mathbf{0}$.
- Let $\mathcal{I}_{\max}$ denote the signal subset that yields the maximum in (19) and thus determines the list radius. Then all the points from $\mathcal{I}_{\max}$ lie on the sphere $\mathscr{S}$ and the remaining points, from $\mathcal{S}_L \setminus \mathcal{I}_{\max}$, lie inside.
- The center $\boldsymbol{\theta}$ of the sphere $\mathscr{S}$ is given by (*cf.* (12))

  $$\boldsymbol{\theta} - \boldsymbol{s}_0 = \frac{1}{2}\boldsymbol{\gamma}_{\mathcal{I}_{\max}}\boldsymbol{\Gamma}_{\mathcal{I}_{\max}}^{-1}\boldsymbol{S}_{\mathcal{I}_{\max}}$$

  where $\boldsymbol{S}_{\mathcal{I}_{\max}}$ is the matrix whose rows are $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, $\boldsymbol{s}_\ell \in \mathcal{I}_{\max}$.
- Since $\mathcal{I}_{\max} \subseteq \mathcal{S}_L$, then

  $$R_L(\boldsymbol{\Gamma}) \leq \widetilde{R}_L(\boldsymbol{\Gamma})$$

  with equality if and only if the list configuration matrix $\boldsymbol{\Gamma}$ fulfills

  $$\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1} \geq \mathbf{0}. \tag{20}$$

  Hence, when determining the list radius of a signal set $\mathcal{S}_L$, the first step is to check whether condition (20) is satisfied. If so, then the circumsphere $\widetilde{\mathscr{S}}$ is the smallest encompassing sphere and $R_L(\boldsymbol{\Gamma}) = \widetilde{R}_L(\boldsymbol{\Gamma}) = \frac{1}{2}\sqrt{\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}^{\mathrm{T}}}$. Otherwise, when at least one component of $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}$ is negative, the sphere $\mathscr{S}$ and its radius $R_L(\boldsymbol{\Gamma})$ are determined by a reduced signal set $\mathcal{I}_{\max} \subset \mathcal{S}_L$.
  In some cases, the negative components of $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}$ indicate which points should be removed from $\mathcal{S}_L$ to obtain the set $\mathcal{I}_{\max}$ (see Examples 2, 3, and 5). However, in general, there is no one-to-one correspondence between the negative components of $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}$ and the set $\mathcal{S}_L \setminus \mathcal{I}_{\max}$ (see Example 6).

Note that the condition (20) is equivalent to $\boldsymbol{\zeta} \geq \mathbf{0}$ (*cf.* (18)), that is, all the coefficients of the decomposition of $\boldsymbol{\rho} - \boldsymbol{s}_0$ in the basis set $\{\boldsymbol{s}_\ell - \boldsymbol{s}_0\}$, $\ell = 1, 2, ..., L$, should be nonnegative. The set of points in space described by the linear combination $\boldsymbol{\zeta}\boldsymbol{S}$, where $\boldsymbol{\zeta} \geq \mathbf{0}$, constitutes an L-dimensional unbounded pyramid, whose vertex is $\boldsymbol{s}_0$ and whose semi-infinite edges run along $\boldsymbol{s}_\ell - \boldsymbol{s}_0$. If the center $\boldsymbol{\rho}$ of the circumsphere $\widetilde{\mathscr{S}}$ lies in this pyramid, then $\widetilde{\mathscr{S}}$ is the smallest sphere that determines the list radius. Otherwise, there exists a smaller sphere whose center is inside this pyramid, and it is found by (19).

When the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$, $\ell = 1, 2, ..., L$, are linearly independent, the inverse of the list configuration matrix is given by

$$\boldsymbol{\Gamma}^{-1} = \frac{1}{\det(\boldsymbol{\Gamma})}\operatorname{adj}(\boldsymbol{\Gamma})$$

where $\operatorname{adj}(\boldsymbol{\Gamma})$ is the adjoint matrix of $\boldsymbol{\Gamma}$. Furthermore, $\det(\boldsymbol{\Gamma}) > 0$. Hence, condition (20) is equivalent to

$$\boldsymbol{\gamma}\operatorname{adj}(\boldsymbol{\Gamma}) \geq \mathbf{0}. \tag{21}$$

In fact, this condition is more general, since it is also applicable for configurations where some of the vectors $\boldsymbol{s}_\ell - \boldsymbol{s}_0$ are linearly dependent. Hence, the expression (19) is easily generalised to hold for *any* signal set $\mathcal{S}_L$ as:

$$R_L(\boldsymbol{\Gamma}) = \max_{\mathcal{I}\,:\,\boldsymbol{\gamma}_\mathcal{I}\operatorname{adj}(\boldsymbol{\Gamma}_\mathcal{I})\geq\mathbf{0}} \left\{ \frac{1}{2}\sqrt{\boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1}\boldsymbol{\gamma}_\mathcal{I}^{\mathrm{T}}} \right\}. \tag{22}$$

Note that for a singular list configuration matrix $\boldsymbol{\Gamma}$, the list radius is not necessarily infinite.

For a given list configuration $\mathcal{S}_L$, a list error with respect to the transmitted signal $\boldsymbol{s}_0$ occurs if the received signal point $\boldsymbol{r}$ falls in the error decision region $\mathcal{D}$, which is the intersection of all pairwise error decision regions $\mathcal{D}_\ell$, $\ell = 1, 2, \ldots, L$, between the signal points $\boldsymbol{s}_\ell$ and $\boldsymbol{s}_0$, that is, $\mathcal{D} = \bigcap_{\ell=1}^{L} \mathcal{D}_\ell$. The point of the region $\mathcal{D}$ that is closest to the signal point $\boldsymbol{s}_0$ is the center $\boldsymbol{\theta}$ of the sphere $\mathscr{S}$. If the list radius (22) is $R_L(\boldsymbol{\Gamma}) = \infty$, the pairwise error decision regions do not intersect, $\bigcap_{\ell=1}^{L} \mathcal{D}_\ell = \emptyset$. For such a list configuration, the probability of a list error for a given transmitted signal $\boldsymbol{s}_0$ is zero since there is no point in space that is simultaneously closer to the L points $\boldsymbol{s}_\ell$, $\ell = 1, 2, ..., L$ than to the point $\boldsymbol{s}_0$ (see Example 4 and Tables I and II for $L = 3$).

The minimum list radius (10) for list size L of a signal constellation $\mathcal{S}$ is obtained as

$$R_{L\,\min} = \min_{\boldsymbol{\Gamma}}\{R_L(\boldsymbol{\Gamma})\}$$

where the minimization is performed over all possible list configuration matrices for a list size L, that is, over all possible signal subsets $\mathcal{S}_L \subseteq \mathcal{S}$. The Euclidean *list distance* for a signal subset $\mathcal{S}_L$ with a list configuration matrix $\boldsymbol{\Gamma}$ is defined as

$$d_{\mathrm{E}L}(\boldsymbol{\Gamma}) \triangleq 2R_L(\boldsymbol{\Gamma}). \tag{23}$$

Thus, from (19) it follows that the squared Euclidean list distance is

$$d_{\mathrm{E}L}^2(\boldsymbol{\Gamma}) = \max_{\mathcal{I}\,:\,\boldsymbol{\gamma}_\mathcal{I}\operatorname{adj}(\boldsymbol{\Gamma}_\mathcal{I})\geq\mathbf{0}} \left\{ \boldsymbol{\gamma}_\mathcal{I}\boldsymbol{\Gamma}_\mathcal{I}^{-1}\boldsymbol{\gamma}_\mathcal{I}^{\mathrm{T}} \right\}. \tag{24}$$

The minimum Euclidean list distance of the signal constellation $\mathcal{S}$ is then

$$d_{\mathrm{E}L\min} = 2R_{L\,\min} = \min_{\boldsymbol{\Gamma}}\{d_{\mathrm{E}L}(\boldsymbol{\Gamma})\}.$$

When the signal constellation $\mathcal{S}$ is a set of bipolar signals corresponding to the binary linear code $\mathcal{C}$, we can also define the Hamming list distance of a code $\mathcal{C}$, for a list size L, as

$$d_{\mathrm{H}L}(\mathbf{\Gamma}) \triangleq \frac{d^2_{\mathrm{E}L}(\mathbf{\Gamma})}{4E_s} = \max_{\mathcal{I}\,:\,\boldsymbol{\gamma}_{\mathrm{H}\mathcal{I}}\,\mathrm{adj}(\mathbf{\Gamma}_{\mathrm{H}\mathcal{I}})\geq \mathbf{0}} \left\{\boldsymbol{\gamma}_{\mathrm{H}\mathcal{I}}\mathbf{\Gamma}^{-1}_{\mathrm{H}\mathcal{I}}\boldsymbol{\gamma}^{\mathrm{T}}_{\mathrm{H}\mathcal{I}}\right\}. \tag{25}$$

The minimum Hamming list distance of a code is then

$$d_{\mathrm{H}L\min} \triangleq \frac{d^2_{\mathrm{E}L\min}}{4E_s} = \frac{R^2_{L\,\min}}{E_s}.$$

The examples of Hamming list distances are shown in Tables I–III. In general, the Hamming list distance is not an integer.

The following examples illustrate the list distance and the list radius for a few signal configurations.

Example 1: Consider the set of $L+1=3$ signal points $\mathcal{S} = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \boldsymbol{s}_2\}$ corresponding to the minimal weight codewords that have minimal pairwise distances $d = d_{\mathrm{E}\min}$, as illustrated in Figure 1. If $\boldsymbol{s}_0$ is the transmitted signal point, then a list error (for list size $L=2$) occurs if the received signal point $\boldsymbol{r}$ falls in the region $\mathcal{D}$ marked in the figure, where both signals $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are closer than $\boldsymbol{s}_0$. If, however, $\boldsymbol{r}$ is outside region $\mathcal{D}$, then $\boldsymbol{s}_0$ will always be one of the two signal points closest to $\boldsymbol{r}$, and hence $\boldsymbol{s}_0$ will be included in the list. The smallest sphere that encompasses the three signal points is the circumsphere, that is, $\mathscr{S} = \widetilde{\mathscr{S}}$, and its center $\boldsymbol{\rho}$ is the center of the equilateral triangle. This is the point of the region $\mathcal{D}$ which is closest to $\boldsymbol{s}_0$. The Euclidean list distance is $d_{\mathrm{E}L} = 2R_L = 2d/\sqrt{3}$, and it is invariant with respect to enumeration of the signal points.

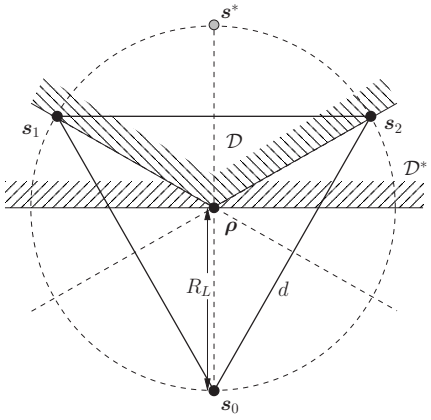

Fig. 1. List configuration and the list error region $\mathcal{D}$ for a list of size $L=2$ with signals at equal pairwise distances. (Region $\mathcal{D}^*$ and point $\boldsymbol{s}^*$ illustrate the bound (42).)

Example 2: Consider now the set $\mathcal{S} = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \boldsymbol{s}_2\}$ of non-bipolar signals shown in Figure 2, where the signal coordinates are $\boldsymbol{s}_0 = (0\ 0)$, $\boldsymbol{s}_1 = (0\ 2)$, and $\boldsymbol{s}_2 = (1\ 3)$. The squared Euclidean distances are $d^2_{\mathrm{E}01} = 4$, $d^2_{\mathrm{E}02} = 3^2 + 1^2 = 10$, and $d^2_{\mathrm{E}12} = 2$. The corresponding Gram matrix is, according to (4),

$$\mathbf{\Gamma} = \begin{pmatrix} 4 & 6 \\ 6 & 10 \end{pmatrix}.$$

Thus we have

$$\boldsymbol{\gamma}\,\mathrm{adj}(\mathbf{\Gamma}) = (\,4\ \ 10\,)\begin{pmatrix} 10 & -6 \\ -6 & 4 \end{pmatrix} = (\,-20\ \ 16\,) \ngeq (\,0\ \ 0\,). \tag{26}$$

The squared radius $\widetilde{R}^2_L$ of the cicumsphere $\widetilde{\mathscr{S}}$ is given by (11):

$$\widetilde{R}^2_L = \frac{1}{4}\,\boldsymbol{\gamma}\mathbf{\Gamma}^{-1}\boldsymbol{\gamma}^{\mathrm{T}} = \frac{1}{4}\,(\,-5\ \ 4\,)\begin{pmatrix} 4 \\ 10 \end{pmatrix} = 5$$

and the coordinates of the center $\boldsymbol{\rho}$ are, according to (12),

$$\boldsymbol{\rho} = \frac{1}{2}\,\boldsymbol{\gamma}\mathbf{\Gamma}^{-1}\boldsymbol{S} = \frac{1}{2}\,(\,-5\ \ 4\,)\begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} = (\,2\ \ 1\,).$$

However, the point of the list error region $\mathcal{D}$ which is nearest to $\boldsymbol{s}_0$ is not $\boldsymbol{\rho}$ but $\boldsymbol{\theta}$, or, equivalently, the sphere $\widetilde{\mathscr{S}}$ is not the smallest sphere which encompasses the three signal points. This is indicated by the negative sign of some of the elements of $\boldsymbol{\gamma}\,\mathrm{adj}(\mathbf{\Gamma})$, as found in (26). Hence, in order to find the list radius and the list distance, we need to reduce the size of the list and find the subset of signal points for which the distance $\boldsymbol{\gamma}_{\mathcal{I}}\mathbf{\Gamma}^{-1}_{\mathcal{I}}\boldsymbol{\gamma}^{\mathrm{T}}_{\mathcal{I}}$ is maximized. In our case, the signal point $\boldsymbol{s}_2$ has larger distance from $\boldsymbol{s}_0$ and thus the signal set is reduced to $\mathcal{I}_{\max} = \{\boldsymbol{s}_0, \boldsymbol{s}_2\}$ for which $\mathbf{\Gamma}_{\mathcal{I}_{\max}} = \boldsymbol{\gamma}_{\mathcal{I}_{\max}} = (\,10\,)$. The

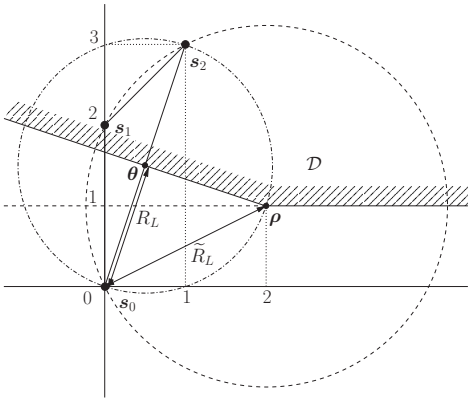

Fig. 2. Configuration of three non-bipolar signals for which the smallest sphere encompassing the points does not coincide with the sphere on which all the points are lying, and thus, $R_L \neq \widetilde{R}_L$.

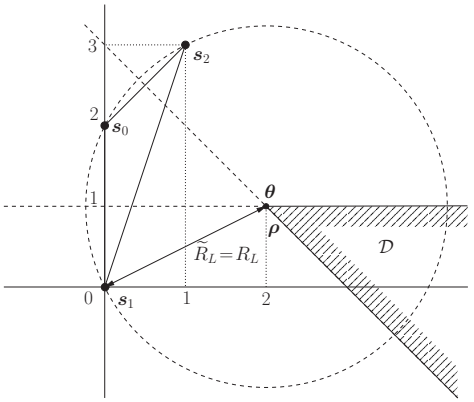

Fig. 3. The same configuration as in Fig. 2, but with changed reference signal point $\boldsymbol{s}_0$: in this case the two spheres coincide and the list error decision region $\mathcal{D}$ has a completely different shape compared to the previous case.

squared list distance is thus

$$d^2_{\mathrm{E}L} = d^2_{\mathrm{E}02} = 10,$$

and the list radius is $R_L = \sqrt{d^2_{\mathrm{E}L}/4} = \sqrt{10}/2 = \sqrt{5/2}$, which is smaller than $\widetilde{R}_L = \sqrt{5}$. The center of the sphere $\mathscr{S}$, of radius R_L, is given by

$$\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{\gamma}_{\mathcal{I}_{\max}}\boldsymbol{\Gamma}^{-1}_{\mathcal{I}_{\max}}\boldsymbol{S}_{\mathcal{I}_{\max}} = \frac{1}{2}\,(1\ \ 3) = \left(\tfrac{1}{2}\quad \tfrac{3}{2}\right).$$

Note that, unlike in Example 1, the list distance for the configuration in Figure 2 is not symmetric with respect to the signal points, that is, it changes if we change the reference signal point. For example, if we exchange signal points $\boldsymbol{s}_1$ and $\boldsymbol{s}_0$, we obtain the configuration as illustrated in Figure 3. The shape of the error decision region $\mathcal{D}$ is changed and, in this case, the two spheres coincide. Thus, the list distance is determined by the radius of the circumsphere, that is, $d^2_{\mathrm{E}L} = 4\widetilde{R}^2_L = 20$.

Example 3: Consider the signal constellation shown in Figure 4, with the signals $\boldsymbol{s}_0 = (0\ 0)$, $\boldsymbol{s}_1 = (0\ 1)$, and $\boldsymbol{s}_2 = (0\ 2)$. The intersection of the pairwise error regions is $\mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2 = \mathcal{D}_2$, as shown in the figure. The squared Euclidean distances are $d^2_{\mathrm{E}01} = 1$, $d^2_{\mathrm{E}02} = 4$, and $d^2_{\mathrm{E}12} = 1$, and the corresponding Gram matrix is

$$\boldsymbol{\Gamma} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

The matrix $\boldsymbol{\Gamma}$ is singular due to linear dependency of the signals. Since $\det(\boldsymbol{\Gamma}) = 0$ it follows from (11) that $\widetilde{R}_L = \infty$, *i.e.*, the circle on which the signal points are lying is a straight line.

Let us now determine the list radius R_L. We have that

$$\boldsymbol{\gamma}\,\mathrm{adj}(\boldsymbol{\Gamma}) = (1\ \ 4)\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} = (-4\ \ 2) \ngeq (0\ \ 0)$$

which indicates that in order to obtain R_L, we need to reduce the list size, as in the previous example. The signal point $\boldsymbol{s}_2$ has larger distance from $\boldsymbol{s}_0$ and thus the signal set is reduced to $\mathcal{I}_{\max} = \{\boldsymbol{s}_0, \boldsymbol{s}_2\}$ for which $\boldsymbol{\Gamma}_{\mathcal{I}_{\max}} = \boldsymbol{\gamma}_{\mathcal{I}_{\max}} = (4)$. The squared list distance is thus

$$d^2_{\mathrm{E}L} = d^2_{\mathrm{E}02} = 4,$$

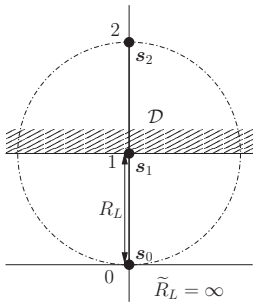

Fig. 4. List configuration with linearly dependent signal points for which $\widetilde{R}_L = \infty$, but the list radius R_L is finite.

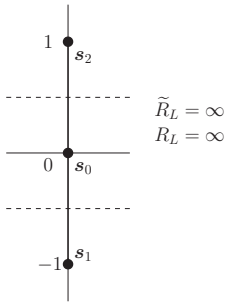

Fig. 5. List configuration with linearly dependent signal points for which $\widetilde{R}_L = R_L = \infty$. The list error probability is 0.

and the list radius is $R_L = \sqrt{d^2_{\mathrm{E}L}/4} = 1$. The center of the sphere $\mathscr{S}$ is the point $\boldsymbol{s}_1$, which is formally obtained by

$$\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{\gamma}_{\mathcal{I}_{\max}}\boldsymbol{\Gamma}^{-1}_{\mathcal{I}_{\max}}\boldsymbol{S}_{\mathcal{I}_{\max}} = \frac{1}{2}\,(0\ 2) = (0\ 1) = \boldsymbol{s}_1.$$

Example 4: Consider the signal constellation shown in Figure 5, with the signals $\boldsymbol{s}_0 = 0$, $\boldsymbol{s}_1 = -1$, and $\boldsymbol{s}_2 = 1$. This configuration corresponds to the one from the previous example, with changed reference point. The pairwise decision regions, indicated by the two dashed lines, do not intersect, that is, $\mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$, which implies that the probability of list error is zero (because the received signal point can never be closer to both $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ than to $\boldsymbol{s}_0$). This corresponds to an infinite list distance as we will now formally verify. The list configuration matrix is

$$\boldsymbol{\Gamma} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

The matrix $\boldsymbol{\Gamma}$ is singular, thus $\det(\boldsymbol{\Gamma}) = 0$ and $\widetilde{R}_L = \infty$. Now we consider the vector $\boldsymbol{\gamma}\,\mathrm{adj}(\boldsymbol{\Gamma})$:

$$\boldsymbol{\gamma}\,\mathrm{adj}(\boldsymbol{\Gamma}) = (1\ \ 1)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = (2\ \ 2) \geq (0\ \ 0).$$

Since the elements of $\boldsymbol{\gamma}\,\mathrm{adj}(\boldsymbol{\Gamma})$ are positive, we conclude that the submatrix $\boldsymbol{\Gamma}_{\mathcal{I}}$ which maximizes the quadratic form $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}^{\mathrm{T}}$ is the matrix $\boldsymbol{\Gamma}$ itself, and therefore $d_{\mathrm{E}L} = 2R_L = 2\widetilde{R}_L = \infty$.

Example 5: Consider the set of $L+1 = 4$ bipolar signal points corresponding to the codewords

$$\begin{aligned} \boldsymbol{v}_0 &= (0\ 0\ 0\ 0\ 0\ 0) \\ \boldsymbol{v}_1 &= (1\ 1\ 1\ 1\ 1\ 0) \\ \boldsymbol{v}_2 &= (0\ 1\ 1\ 1\ 1\ 1) \\ \boldsymbol{v}_3 &= (0\ 1\ 1\ 0\ 0\ 0). \end{aligned}$$

The pairwise Hamming distances are $d_{\mathrm{H}01} = d_{\mathrm{H}02} = 5$, $d_{\mathrm{H}03} = 2$, $d_{\mathrm{H}12} = 2$, $d_{\mathrm{H}13} = d_{\mathrm{H}23} = 3$, which yields the list configuration matrix

$$\boldsymbol{\Gamma} = 4E_s\begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

The $L = 3$ dimensional circumsphere has the squared radius $\widetilde{R}^2_L = \frac{1}{4}\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}^{\mathrm{T}} = 1.4 \cdot 4E_s$. However,

$$\boldsymbol{\gamma}\,\mathrm{adj}(\boldsymbol{\Gamma}) = (6\ \ 6\ \ -2) \ngeq (0\ \ 0\ \ 0)$$

which implies that the list distance $d_{\mathrm{E}L}$ is determined by a reduced signal set. In our case, we need to remove the sequence $\boldsymbol{v}_3$ in order to maximize the quadratic form (24) and, thus, for $\mathcal{I}_{\max} = \{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2\}$ we obtain

$$d^2_{\mathrm{E}L} = \boldsymbol{\gamma}_{\mathcal{I}_{\max}}\boldsymbol{\Gamma}^{-1}_{\mathcal{I}_{\max}}\boldsymbol{\gamma}^{\mathrm{T}}_{\mathcal{I}_{\max}} = 5.56 \cdot 4E_s$$

and the squared list radius is $R^2_L = d^2_{\mathrm{E}L}/4 = 1.389 \cdot 4E_s$. A sphere of the radius R_L contains the signal points from $\mathcal{I}_{\max}$ on its surface, while $\boldsymbol{v}_3$ lies inside.

Example 6: Consider a set of $L+1=4$ points in the three-dimensional space, $\mathcal{S}_L=\{\boldsymbol{s}_0,\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3\}$, where the reference point is the origin, $\boldsymbol{s}_0=(0\ 0\ 0)$, and the remaining points are

$$\begin{aligned}\boldsymbol{s}_1 &= (4\ 3\ 4)\\ \boldsymbol{s}_2 &= (4\ 3\ 2)\\ \boldsymbol{s}_3 &= (3\ 1\ 1).\end{aligned}$$

The list configuration matrix is

$$\boldsymbol{\Gamma}=\begin{pmatrix}41 & 33 & 19\\ 33 & 29 & 17\\ 19 & 17 & 11\end{pmatrix}.$$

The circumsphere $\widetilde{\mathscr{S}}$ of the set $\mathcal{S}_L$ has the radius $\widetilde{R}_L=4.32$, and it is illustrated in Figure 6. Since

$$\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}=(1.8\ 2\ -5.2)\not\geq(0\ 0\ 0)$$

the list radius is determined by a reduced signal set. Using (22), we find that the smallest sphere $\mathscr{S}$ encompassing the signal points contains only signals $\mathcal{I}_{\max}=\{\boldsymbol{s}_0,\boldsymbol{s}_1\}$ on its surface, while the points $\mathcal{S}_L\setminus\mathcal{I}_{\max}=\{\boldsymbol{s}_2,\boldsymbol{s}_3\}$ lie inside. Note that only the third element of $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}$ is negative; however, both $\boldsymbol{s}_2$ and $\boldsymbol{s}_3$ are inside $\mathscr{S}$. The sphere $\mathscr{S}$ is illustrated in Figure 6. Points $\boldsymbol{s}_0$ and $\boldsymbol{s}_1$ are visible on the intersections of the two spheres. The corresponding list radius is

$$R_L=\frac{1}{2}\sqrt{41}\approx 3.2.$$

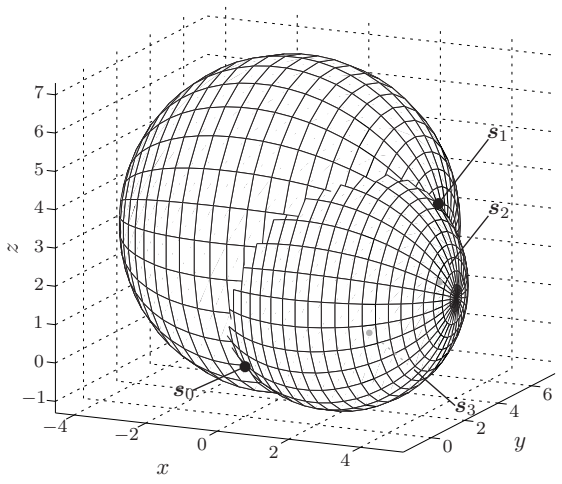

Fig. 6. The circumsphere $\widetilde{\mathscr{S}}$ and the smallest encompassing sphere $\mathscr{S}$ for the signal set from Example 6.

C. Properties of the List Configuration Matrix

Hereinafter, we consider sets of linearly independent bipolar code signals $\boldsymbol{s}_\ell$, $\ell=1,2,\ldots,L$, for which the matrix $\boldsymbol{\Gamma}$ fulfills the condition $\boldsymbol{\gamma}\boldsymbol{\Gamma}^{-1}\geq\mathbf{0}$. The following theorem establishes a connection between the Hamming list distance $d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}})$ and some properties of the matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$.

Theorem 2: Let $\boldsymbol{\Gamma}_{\mathrm{H}}=\{\gamma_{\mathrm{H}ij}\}$, $i,j=1,2,\ldots,L$, be the normalized Gram matrix (with entries given by (8)) of linearly independent bipolar signal vectors $\boldsymbol{s}_i-\boldsymbol{s}_0$ of the binary block code with the minimum Hamming distance $d_{\mathrm{H}\min}$. Then the quadratic form $d_{\mathrm{H}L}(\boldsymbol{\Gamma})=\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}^{-1}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}}$ has the following properties:

1) All $\gamma_{\mathrm{H}ij}$ are integers, and $\gamma_{\mathrm{H}ii}\geq\gamma_{\mathrm{H}ij}\geq 0$.
2) $\boldsymbol{\Gamma}_{\mathrm{H}}$ is positive definite.
3) If $\lambda_{\max}$ is the maximal eigenvalue of $\boldsymbol{\Gamma}_{\mathrm{H}}$, then the Hamming list distance satisfies
$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}})\geq\frac{||\boldsymbol{\gamma}_{\mathrm{H}}||^4}{\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}}}\geq||\boldsymbol{\gamma}_{\mathrm{H}}||^2/\lambda_{\max}.$$
4) For any binary code with the minimum Hamming distance $d_{\mathrm{H}\min}$, the minimum Hamming list distance is
$$d_{\mathrm{H}L\min}\geq\frac{2L}{L+1}d_{\mathrm{H}\min}$$
where equality is achieved for even $d_{\mathrm{H}\min}$ with a matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ whose main-diagonal and off-diagonal elements are $\gamma_{\mathrm{H}ii}=d_{\mathrm{H}\min}$ and $\gamma_{\mathrm{H}ij}=d_{\mathrm{H}\min}/2$, respectively, if such a matrix exists.
5) For any binary code with odd $d_{\mathrm{H}\min}$, the minimum Hamming list distance is
$$d_{\mathrm{H}L\min}\geq\frac{2L}{L+1}d_{\mathrm{H}\min}+\frac{L-1}{L+1}$$
where the equality is achieved for odd list size L and a matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ with main diagonal elements
$$\gamma_{\mathrm{H}ii}=\begin{cases}d_{\mathrm{H}\min}, & 1\leq i\leq\frac{L+1}{2}\\ d_{\mathrm{H}\min}+1, & \frac{L+1}{2}+1\leq i\leq L\end{cases}$$
and off-diagonal elements
$$\gamma_{\mathrm{H}ij}=\begin{cases}(d_{\mathrm{H}\min}-1)/2, & \text{if } \gamma_{\mathrm{H}ii}=\gamma_{\mathrm{H}jj}=d_{\mathrm{H}\min}\\ (d_{\mathrm{H}\min}+1)/2, & \text{otherwise}\end{cases}$$
if such a matrix exists.

Remark: From Statement 4 of Theorem 2, it follows that the ratio between the minimum Hamming list distance $d_{\mathrm{H}L\min}$ and the minimum distance of the code $d_{\mathrm{H}\min}$ (when $d_{\mathrm{H}\min}$ is even) is
$$\frac{d_{\mathrm{H}L\min}}{d_{\mathrm{H}\min}}\geq\frac{2L}{L+1}.$$
The ratio on the right-hand side of the above inequality was derived in [2] using simplex geometry and defined as the *asymptotic list decoding gain* (over ML decoding).

Proof:

1) The entries of the matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ are given by (8). Since the Hamming distance satisfies the triangle inequality, we immediately obtain that $\gamma_{\mathrm{H}ij}=(d_{\mathrm{H}0i}+d_{\mathrm{H}0j}-d_{\mathrm{H}ij})/2\geq 0$. Furthermore, we have
$$\begin{aligned}\gamma_{\mathrm{H}ij}-\gamma_{\mathrm{H}ii} &= \frac{d_{\mathrm{H}0i}+d_{\mathrm{H}0j}-d_{\mathrm{H}ij}}{2}-d_{\mathrm{H}0i}\\ &= \frac{d_{\mathrm{H}0j}-(d_{\mathrm{H}0i}+d_{\mathrm{H}ij})}{2}\leq 0\end{aligned}$$
which yields $\gamma_{\mathrm{H}ii}\geq\gamma_{\mathrm{H}ij}\geq 0$. In order to prove that the entries of the matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ are integers, it suffices to verify that $d_{\mathrm{H}0i}+d_{\mathrm{H}0j}-d_{\mathrm{H}ij}$ is always an even number. This

follows directly from the fact that when two codewords have weights $d_{\mathrm{H}0i}$ and $d_{\mathrm{H}0j}$ of the same parity (both odd or both even), then their pairwise distance $d_{\mathrm{H}ij}$ is always even; while for codeword weights of opposite parity, the pairwise distance is an odd number.

2) The matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ is a Gram matrix of linearly independent vectors $(\boldsymbol{s}_i - \boldsymbol{s}_0)$, normalized by $4E_s$. Hence, $\boldsymbol{\Gamma}_{\mathrm{H}}$ is positive definite (*cf.* Theorem 7.2.10 in [29]).
3) According to the Kantorovich inequality [29], for every positive definite symmetric matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ and any non-negative row vector $\boldsymbol{x} \geq \boldsymbol{0}$ we have

$$\frac{4\lambda_{\min}\lambda_{\max}}{(\lambda_{\min}+\lambda_{\max})^2} \leq \frac{(\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}})^2}{(\boldsymbol{x}\boldsymbol{\Gamma}_{\mathrm{H}}\boldsymbol{x}^{\mathrm{T}})(\boldsymbol{x}\boldsymbol{\Gamma}_{\mathrm{H}}^{-1}\boldsymbol{x}^{\mathrm{T}})} \leq 1 \quad (27)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and the largest eigenvalue of $\boldsymbol{\Gamma}_{\mathrm{H}}$, respectively. By applying the right side of the Kantorovich inequality (27) with $\boldsymbol{x} = \boldsymbol{\gamma}_{\mathrm{H}}$ we obtain

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) = \boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}^{-1}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}} \geq \frac{||\boldsymbol{\gamma}_{\mathrm{H}}||^4}{\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}}}.$$

Furthermore, according to Theorem 4.2.2 in [29], $\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}} \leq ||\boldsymbol{\gamma}_{\mathrm{H}}||^2\lambda_{\max}$. Thus we obtain

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) \geq \frac{||\boldsymbol{\gamma}_{\mathrm{H}}||^4}{\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}}} \geq \frac{||\boldsymbol{\gamma}_{\mathrm{H}}||^2}{\lambda_{\max}}.$$

4) The matrix $\boldsymbol{\Gamma}_{\mathrm{H}}$ can be decomposed as follows

$$\boldsymbol{\Gamma}_{\mathrm{H}} = \boldsymbol{D}\boldsymbol{V}\boldsymbol{D}^{\mathrm{T}}$$

where $\boldsymbol{D}$ is the diagonal matrix with entries $\sqrt{\gamma_{\mathrm{H}ii}}$ and $\boldsymbol{V} = \left\{w_{\mathrm{H}ij}/\sqrt{w_{\mathrm{H}ii}w_{\mathrm{H}jj}}\right\}$, $i,j = 1,2,\ldots,L$. Then the Hamming list distance is

$$\begin{aligned} d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) = \boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{\Gamma}_{\mathrm{H}}^{-1}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}} &= \underbrace{\left(\boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{D}^{-1}\right)}_{\triangleq \boldsymbol{v}}\boldsymbol{V}^{-1}\left(\boldsymbol{D}^{-1}\boldsymbol{\gamma}_{\mathrm{H}}^{\mathrm{T}}\right) \\ &= \boldsymbol{v}\boldsymbol{V}^{-1}\boldsymbol{v}^{\mathrm{T}} \end{aligned}$$

where $\boldsymbol{v} = \boldsymbol{\gamma}_{\mathrm{H}}\boldsymbol{D}^{-1} = \left(\sqrt{w_{\mathrm{H}11}}\ \sqrt{w_{\mathrm{H}22}}\ \cdots\ \sqrt{w_{\mathrm{H}LL}}\right)$. Since $\boldsymbol{V}$ is positive definite and $\boldsymbol{v} > \boldsymbol{0}$, we apply the Kantorovich inequality (27) and obtain

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) = \boldsymbol{v}\boldsymbol{V}^{-1}\boldsymbol{v}^{\mathrm{T}} \geq \frac{||\boldsymbol{v}||^4}{\boldsymbol{v}\boldsymbol{V}\boldsymbol{v}^{\mathrm{T}}}.$$

Since

$$||\boldsymbol{v}||^4 = \left(\sum_{i=1}^{L}\gamma_{\mathrm{H}ii}\right)^2 = \left(\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})\right)^2$$

and

$$\begin{aligned} \boldsymbol{v}\boldsymbol{V}\boldsymbol{v}^{\mathrm{T}} &= \sum_{i=1}^{L}\sum_{j=1}^{L}\sqrt{\gamma_{\mathrm{H}ii}}\frac{\gamma_{\mathrm{H}ij}}{\sqrt{\gamma_{\mathrm{H}ii}\gamma_{\mathrm{H}jj}}}\sqrt{\gamma_{\mathrm{H}jj}} \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L}\gamma_{\mathrm{H}ij} = \sum_{i=1}^{L}\sum_{j=1}^{L}\frac{d_{\mathrm{H}0i}+d_{\mathrm{H}0j}-d_{\mathrm{H}ij}}{2} \\ &= \frac{1}{2}\left(2\sum_{j=1}^{L}\sum_{i=1}^{L}d_{\mathrm{H}0i} - \sum_{i=1}^{L}\sum_{j=1}^{L}d_{\mathrm{H}ij}\right) \\ &= \frac{1}{2}\left(2L\,\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) - \sum_{i=1}^{L}\sum_{j=1}^{L}d_{\mathrm{H}ij}\right) \\ &= L\,\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) - \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}d_{\mathrm{H}ij} \end{aligned}$$

we obtain

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) \geq \frac{\left(\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})\right)^2}{L\,\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) - \frac{1}{2}\sum\limits_{i=1}^{L}\sum\limits_{j=1}^{L}d_{\mathrm{H}ij}}. \quad (28)$$

Since the pairwise distance between any two codewords satisfies $d_{\mathrm{H}ij} \geq d_{\mathrm{H}\min}$, for $i \neq j$, and $d_{\mathrm{H}ij} = 0$, for $i = j$, then we have

$$\sum_{i=1}^{L}\sum_{j=1}^{L}d_{\mathrm{H}ij} \geq L(L-1)d_{\mathrm{H}\min}$$

which, combined with (28), yields

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) \geq \frac{\left(\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})\right)^2}{L\,\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) - \frac{1}{2}L(L-1)d_{\mathrm{H}\min}}. \quad (29)$$

By taking the derivative of the right-hand side of (29) with respect to $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})$ we find that its minimum is achieved for $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) = (L-1)d_{\mathrm{H}\min}$. On the other hand, all the diagonal entries of $\boldsymbol{\Gamma}_{\mathrm{H}}$ are $\gamma_{\mathrm{H}ii} \geq d_{\mathrm{H}\min}$, which implies that $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) \geq Ld_{\mathrm{H}\min}$ must hold. Clearly, in this range of $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})$, (29) is a monotonically increasing function of $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}})$; hence, its minimum value is obtained for $\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{H}}) = Ld_{\mathrm{H}\min}$ and we obtain the following bound

$$d_{\mathrm{H}L}(\boldsymbol{\Gamma}_{\mathrm{H}}) \geq \frac{2L}{L+1}d_{\mathrm{H}\min}. \quad (30)$$

Equality in (30) is achieved for a code with even minimum Hamming distance $d_{\mathrm{H}\min}$ and a list configuration matrix of the form

$$\boldsymbol{\Gamma}_{\mathrm{H}} = \begin{pmatrix} d_{\mathrm{H}\min} & \frac{d_{\mathrm{H}\min}}{2} & \cdots & \frac{d_{\mathrm{H}\min}}{2} \\ \frac{d_{\mathrm{H}\min}}{2} & d_{\mathrm{H}\min} & \cdots & \frac{d_{\mathrm{H}\min}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d_{\mathrm{H}\min}}{2} & \frac{d_{\mathrm{H}\min}}{2} & \cdots & d_{\mathrm{H}\min} \end{pmatrix}. \quad (31)$$

We call this matrix the worst-case list configuration matrix; it corresponds to the list of L codewords with minimum weights and at minimal possible pairwise distances, which are, for even $d_{\mathrm{H}\min}$, all equal to $d_{\mathrm{H}\min}$ (*cf.* Tables II and III). It is easy to verify that matrix (31) satisfies (30) with equality.

5) If the Hamming weights of two codewords have the same parity (both odd or both even), the Hamming distance between them is always even. Hence, if a code has odd minimum distance d_{Hmin}, the pairwise distances between codewords of weight d_{Hmin} must be even, that is, $d_{\mathrm{H}ij} \geq d_{\mathrm{Hmin}}+1$. Then, if d_{Hmin} is odd, we conclude that the worst-case list configuration matrix (31) cannot be constructed and the bound (30) is never tight.
To obtain the worst-case $\mathbf{\Gamma}_{\mathrm{H}}$ and the corresponding minimum list distance when d_{Hmin} is odd, assume that in a list of L codewords there are $m \geq 0$ codewords of odd weight and $L-m$ codewords of even weight. The pairwise distances between the codewords of same-parity weights are $d_{\mathrm{H}ij} \geq d_{\mathrm{Hmin}}+1$, while for opposite-parity pairs $d_{\mathrm{H}ij} \geq d_{\mathrm{Hmin}}$, which yields

$$\sum_{i=1}^{L}\sum_{j=1}^{L} d_{\mathrm{H}ij} \geq L(L-1)d_{\mathrm{Hmin}}+m(m-1)+(L-m)(L-m-1). \tag{32}$$

Following the same procedure as in the previous case, we insert (32) into (28) and, by taking the derivative of the right-hand side of (28) with respect to $\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{H}})$, we conclude that the bound (28) is a monotonically increasing function of $\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{H}})$ for

$$\begin{aligned}\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{H}}) &\geq \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L} d_{\mathrm{H}ij} \\ &\geq (L-1)d_{\mathrm{Hmin}} + \frac{2m^2}{L} - 2m + L - 1. \end{aligned}\tag{33}$$

On the other hand, since the m odd-weight codewords from the list have weights $d_{\mathrm{H}0i} \geq d_{\mathrm{Hmin}}$, while the remaining $(L-m)$ even-weight codewords have $d_{\mathrm{H}0i} \geq d_{\mathrm{Hmin}}+1$, then $\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{H}})$ has to fulfill

$$\begin{aligned}\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{H}}) &\geq md_{\mathrm{Hmin}} + (L-m)(d_{\mathrm{Hmin}}+1) \\ &= L(d_{\mathrm{Hmin}}+1) - m. \end{aligned}\tag{34}$$

The right-hand side of (34) is always larger than the right-hand side of (33); hence, we conclude that the right-hand side of (34) minimizes the bound (28) on the list distance. Thus we obtain

$$d_{\mathrm{H}L}(\mathbf{\Gamma}_{\mathrm{H}}) \geq \frac{2(L(d_{\mathrm{Hmin}}+1)-m)^2}{L(L+1)(d_{\mathrm{Hmin}}+1)-2m^2} \tag{35}$$

which holds for all list configuration matrices $\mathbf{\Gamma}_{\mathrm{H}}$. By taking the derivative of (35) with respect to m we can conclude that, assuming odd L, the minimum of the bound (35) is achieved for $m=(L+1)/2$, which yields

$$d_{\mathrm{H}L}(\mathbf{\Gamma}_{\mathrm{H}}) \geq \frac{2L}{L+1}d_{\mathrm{Hmin}} + \frac{L-1}{L+1}. \tag{36}$$

Equality is achieved for the worst-case list configuration matrix with $m=(L+1)/2$ diagonal elements equal to $\gamma_{\mathrm{H}ii} = d_{\mathrm{Hmin}}$ and $L-m$ elements $\gamma_{\mathrm{H}ii} = d_{\mathrm{Hmin}}+1$, and with the off-diagonal elements equal to $\gamma_{\mathrm{H}ij} = (d_{\mathrm{Hmin}}-1)/2$ if $\gamma_{\mathrm{H}ii} = \gamma_{\mathrm{H}jj} = d_{\mathrm{Hmin}}$, and $\gamma_{\mathrm{H}ij} = (d_{\mathrm{Hmin}}+1)/2$ otherwise. If the list size L is even, the worst-case matrix is obtained in the same way, with $m = \lceil (L+1)/2 \rceil$; however, in this case, the lower bound (35) on the list distance is not tight.

■

D. Center of Mass and Average Radius of a List

For a given set $\mathcal{S}_L = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$ of $L+1$ signal points, the *center of mass* is located in the point $\bar{\boldsymbol{s}}$ given by

$$\bar{\boldsymbol{s}} = \frac{1}{L+1}\sum_{j=0}^{L} \boldsymbol{s}_j. \tag{37}$$

The *average squared radius* $R^2_{L\mathrm{av}}$, introduced in [12], of the signal set $\mathcal{S}_L$ is the average squared Euclidean distance of the signal points $\boldsymbol{s}_i \in \mathcal{S}_L$ from the center of mass $\bar{\boldsymbol{s}}$, that is,

$$R^2_{L\mathrm{av}} = \frac{1}{L+1}\sum_{i=0}^{L} ||\boldsymbol{s}_i - \bar{\boldsymbol{s}}||^2 = \frac{1}{L+1}\sum_{i=0}^{L} d^2_{\mathrm{E}}(\boldsymbol{s}_i, \bar{\boldsymbol{s}}). \tag{38}$$

The average squared distance between the signal points from a set $\mathcal{S}_L$ and a given reference point is often referred to as the *moment of inertia* of $\mathcal{S}_L$, *cf.* [16]. Clearly, the moment of inertia is smallest when the reference point is the center of mass $\bar{\boldsymbol{s}}$, and then it equals the average squared radius $R^2_{L\mathrm{av}}$.

From the above definitions it follows that the average radius is never larger than the list radius for the given list $\mathcal{S}_L$ with the configuration matrix $\mathbf{\Gamma}$, that is,

$$R^2_L(\mathbf{\Gamma}) \geq R^2_{L\mathrm{av}}. \tag{39}$$

By substituting (37) into (38) we obtain that the average squared radius can also be written as

$$\begin{aligned} R^2_{L\mathrm{av}} &= \frac{1}{L+1}\sum_{i=0}^{L}\Big\|\boldsymbol{s}_i - \frac{1}{L+1}\sum_{j=0}^{L}\boldsymbol{s}_j\Big\|^2 \\ &= \frac{1}{(L+1)^2}\sum_{j=0}^{L}\sum_{i=0}^{L}||\boldsymbol{s}_i||^2 - \frac{2}{(L+1)^2}\sum_{i=0}^{L}\sum_{j=0}^{L}\langle \boldsymbol{s}_i, \boldsymbol{s}_j\rangle \\ &\qquad + \frac{1}{(L+1)^3}\sum_{i=0}^{L}\Big\|\sum_{j=0}^{L}\boldsymbol{s}_j\Big\|^2 \\ &= \frac{1}{(L+1)^2}\sum_{i=0}^{L}\sum_{j=0}^{L}\Big(||\boldsymbol{s}_i||^2 - \langle \boldsymbol{s}_i, \boldsymbol{s}_j\rangle\Big) \\ &= \frac{1}{2(L+1)^2}\sum_{i=0}^{L}\sum_{j=0}^{L}||\boldsymbol{s}_i - \boldsymbol{s}_j||^2 \\ &= \frac{1}{2(L+1)^2}\sum_{i=0}^{L}\sum_{j=0}^{L} d^2_{\mathrm{E}}(\boldsymbol{s}_i, \boldsymbol{s}_j) \\ &\geq \frac{1}{2(L+1)^2}L(L+1)d^2_{\mathrm{Emin}} = \frac{L}{2(L+1)}d^2_{\mathrm{Emin}} \end{aligned}$$

which, combined with (39) yields

$$R^2_L(\mathbf{\Gamma}) \geq R^2_{L\mathrm{av}} \geq \frac{L}{2(L+1)}d^2_{\mathrm{Emin}} \tag{40}$$

with equality when all $(L+1)$ points have minimum pairwise distances d_{Emin}, that is, when they form an L-dimensional

regular simplex. In this case, the center of mass of $\mathcal{S}_L$ coincides with the center of the circumsphere of $\mathcal{S}_L$, that is, the minimum average radius is equal to the minimum list radius,

$$R^2_{L\,\mathrm{av\,min}} = R^2_{L\,\mathrm{min}} = \frac{Ld^2_{\mathrm{Emin}}}{2(L+1)}.$$

When the signal vectors $\boldsymbol{s}_i \in \mathcal{S}_L$ are bipolar sequences of a binary block code $\mathcal{C}$ with minimum Hamming distance d_{Hmin}, then (7) and (25) hold; hence (40) yields the bound

$$d_{\mathrm{HL}\,\mathrm{min}} \geq d_{\mathrm{HL}}(\boldsymbol{\Gamma}) \geq \frac{2L}{(L+1)} d_{\mathrm{Hmin}}$$

which coincides with Statement 4 of Theorem 2. Statement 5 of Theorem 2 can be proved similarly, via the average radius, taking into account the parity of pairwise distances.

The average radius and the moment of inertia of a list were used in [12] and [16] (*cf.* also [30]) for deriving asymptotic bounds on the code rates and list error performance.

III. Upper Bound on the List Error Probability for a Given List

Using properties of the list configuration matrix $\boldsymbol{\Gamma}$ we can upper-bound the list error probability $\Pr(\boldsymbol{t} \geq \boldsymbol{\gamma}/2)$ in (5). In [27] the following Chernoff-type bound was proved

$$P_{\mathrm{e}L}(\boldsymbol{\Gamma}) = \Pr(\boldsymbol{t} \geq \boldsymbol{\gamma}/2) \leq \exp\Big(-d^2_{\mathrm{E}L}(\boldsymbol{\Gamma})/(4N_0)\Big). \quad (41)$$

From (11), (13), and (23) it immediately follows that, for a given list configuration $\boldsymbol{\Gamma}$, the probability of list error is not larger than the probability that the noise component along $\boldsymbol{\rho} - \boldsymbol{s}_0$ is larger than the radius $\widetilde{R}_L$, that is,

$$P_{\mathrm{e}L}(\boldsymbol{\Gamma}) \leq \Pr\left(\nu \geq \widetilde{R}_L(\boldsymbol{\Gamma})\right) \leq \Pr\left(\nu \geq R_L(\boldsymbol{\Gamma})\right)$$

where

$$\nu = \frac{\langle \boldsymbol{n}, \boldsymbol{\rho} - \boldsymbol{s}_0 \rangle}{||\boldsymbol{\rho} - \boldsymbol{s}_0||}$$

is the noise component along $\boldsymbol{\rho} - \boldsymbol{s}_0$. The above inequalities are met with equality for $L = 1$. Since ν is a zero-mean Gaussian random variable with variance $N_0/2$ we obtain

$$P_{\mathrm{e}L}(\boldsymbol{\Gamma}) \leq Q\left(\frac{R_L(\boldsymbol{\Gamma})}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{2d_{\mathrm{HL}}(\boldsymbol{\Gamma})\frac{E_s}{N_0}}\right) \quad (42)$$

where $Q(x) = 1/\sqrt{2\pi}\int_x^\infty \exp(-y^2/2)\mathrm{d}y$. It is easy to see that the bound (42) is tighter than (41). Figure 1 illustrates the worst-case list configuration for $L = 2$. In this case, the bound (42) corresponds to the probability that the received signal falls into the decision region $\mathcal{D}^*$, which is the upper half-plane containing the sphere center $\boldsymbol{\rho}$. Note that if $\boldsymbol{s}^*$ denotes the virtual "average signal point" ("average" of the set $\{\boldsymbol{s}_1, \boldsymbol{s}_2\}$) as shown in Figure 1, then the half-plane $\mathcal{D}^*$ corresponds to the error region for the pairwise error event between $\boldsymbol{s}_0$ and $\boldsymbol{s}^*$.

Now we will derive a new upper bound on $P_{\mathrm{e}L}(\boldsymbol{\Gamma})$ for the worst-case list configuration, which is tighter than the bound (42). We follow an approach similar to the one described in [28]. First we orthogonalize the noise components and then estimate the variances and integration limits for the system of the transformed noise components.

The derivations are based on the following two lemmas, which correspond to the worst-case matrix $\boldsymbol{\Gamma}$ for even and odd minimum distance, respectively.

Hereinafter, we will use notations $\mathbf{1}_{(m)}$ and $\mathbf{0}_{(n)}$ to denote vectors containing m ones and n zeros, respectively. Thus, for example, vector $(a\ a\ a\ b\ 0\ 0)$ can be written as $(a\mathbf{1}_{(3)}\ b\ \mathbf{0}_{(2)}) = a(\mathbf{1}_{(3)}\ b/a\ \mathbf{0}_{(2)})$.

Lemma 1: Let $\boldsymbol{K}$ be an $L \times L$ matrix with the following structure

$$\boldsymbol{K} = \beta\{k_{ij}\}, \quad k_{ij} = \begin{cases} 1, & i = j \\ \kappa, & i \neq j \end{cases}, \quad i, j = 1, 2, \ldots, L \quad (43)$$

where β and κ are arbitrary constants. Then its eigenvalues are

$$\lambda_1 = \beta(1 + \kappa(L-1)) \quad (44)$$
$$\lambda_\ell = \beta(1-\kappa), \quad \ell = 2, 3, \ldots, L \quad (45)$$

with the corresponding eigenvectors

$$\boldsymbol{x}_1 = (1\ 1\ \ldots\ 1) = \mathbf{1}_{(L)} \quad (46)$$
$$\boldsymbol{x}_\ell = \frac{1}{\ell - 1}\left(-\mathbf{1}_{(\ell-1)}\ \ell{-}1\ \mathbf{0}_{(L-\ell)}\right),\ \ell = 2, 3, \ldots, L. \quad (47)$$

Lemma 2: Let $\boldsymbol{K}$ be an $L \times L$ matrix of the following structure

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{C} \end{pmatrix} \quad (48)$$

where $\boldsymbol{A}$ is an $m \times m$ matrix of the form

$$\boldsymbol{A} = \{a_{ij}\}, \quad a_{ij} = \begin{cases} a_0, & i = j \\ a, & i \neq j \end{cases}, \quad i, j = 1, 2, \ldots, m \quad (49)$$

$\boldsymbol{C}$ is an $n \times n$ matrix, with $n = L - m$, of the form

$$\boldsymbol{C} = \{c_{ij}\}, \quad c_{ij} = \begin{cases} c_0, & i = j \\ c, & i \neq j \end{cases}, \quad i, j = 1, 2, \ldots, n \quad (50)$$

and $\boldsymbol{B}$ is an $m \times n$ matrix whose elements are all equal to b:

$$\boldsymbol{B} = \{b_{ij}\}, \quad b_{ij} = b, \quad i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n \quad (51)$$

where a_0, a, c_0, c, and b are arbitrary constant values.

Then the eigenvalues of $\boldsymbol{K}$ are

$$\xi_\ell = a_0 - a, \quad \ell = 1, 2, \ldots, m-1 \quad (52)$$
$$\xi_\ell = c_0 - c, \quad \ell = m, m+1, \ldots, L-2 \quad (53)$$
$$\xi_{L-1} = \frac{(\lambda_{\mathrm{A}1} + \lambda_{\mathrm{C}1})}{2} - \frac{1}{2}\sqrt{(\lambda_{\mathrm{A}1} - \lambda_{\mathrm{C}1})^2 + 4b^2m^2n} \quad (54)$$
$$\xi_L = \frac{(\lambda_{\mathrm{A}1} + \lambda_{\mathrm{C}1})}{2} + \frac{1}{2}\sqrt{(\lambda_{\mathrm{A}1} - \lambda_{\mathrm{C}1})^2 + 4b^2m^2n} \quad (55)$$

where $\lambda_{\mathrm{A}1}$ and $\lambda_{\mathrm{C}1}$ are the dominant eigenvalues of the matrices $\boldsymbol{A}$ and $\boldsymbol{C}$, respectively, that is,

$$\lambda_{\mathrm{A}1} = a_0 + a(m-1) \quad (56)$$
$$\lambda_{\mathrm{C}1} = c_0 + c(n-1). \quad (57)$$

Furthermore, the corresponding eigenvectors of the matrix $\boldsymbol{K}$ are

$$\begin{aligned}
\boldsymbol{x}_\ell &= \frac{1}{\ell}\left(-\mathbf{1}_{(\ell)}\,\ell\,\mathbf{0}_{(L-\ell-1)}\right), \quad \ell = 1,2,\ldots,m-1\\
\boldsymbol{x}_\ell &= \frac{1}{\ell-m+1}\left(\mathbf{0}_{(m)}\ -\mathbf{1}_{(\ell-m+1)}\ \ell{-}m{+}1\ \ \mathbf{0}_{(L-\ell-2)}\right),\\
&\qquad\qquad \ell = m,\, m+1,\ldots,L-2\\
\boldsymbol{x}_{L-1} &= \left(bn{\cdot}\mathbf{1}_{(m)}\ \ (\xi_{L-1}{-}\lambda_{\mathrm{A}1}){\cdot}\mathbf{1}_{(n)}\right)\\
\boldsymbol{x}_L &= \left(bn{\cdot}\mathbf{1}_{(m)}\ \ (\xi_L{-}\lambda_{\mathrm{A}1}){\cdot}\mathbf{1}_{(n)}\right).
\end{aligned}$$

The proofs of Lemmas 1 and 2 are given in Appendix. Now we are ready to state the following two theorems which we use to obtain upper-bounds on the list error probability $P_{\mathrm{e}L}(\boldsymbol{\Gamma})$, for the worst-case list configuration, for even and odd minimum distance, respectively.

Theorem 3: Let $\boldsymbol{t}$ be a Gaussian random vector of length L with zero mean and covariance matrix $\boldsymbol{K}$ given by (43) from Lemma 1, and let $\boldsymbol{\alpha} = \alpha\mathbf{1}_{(L)}$ be a vector of L constant values α. Then the probability $\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha})$ can be upper-bounded by

$$\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha}) \leq \int_{\frac{\alpha L}{\sqrt{\sigma_1}}}^{\infty} f(y) \prod_{\ell=2}^{L} \left(\int_{v_\ell(y)}^{u_\ell(y)} f(x)\mathrm{d}x \right) \mathrm{d}y \qquad (58)$$

with equality for $L \leq 2$. The integration limits are given by

$$u_\ell(y) = \frac{y\sqrt{\sigma_1} - \alpha L}{\sqrt{\sigma_\ell}}, \quad v_\ell(y) = -\frac{u_\ell(y)}{\ell - 1}$$

where $\sigma_1 = L\lambda_1$, with λ_1 given by (44), and $\sigma_\ell = \lambda_\ell \ell/(\ell{-}1)$, $\ell = 2,3,\ldots,L$, with λ_ℓ given by (45). Hereinafter, $f(x)$ and $f(y)$ denote the Gaussian $\mathcal{N}(0,1)$ probability density function.

Theorem 4: Let $\boldsymbol{t}$ be a Gaussian random vector of length L with zero mean and covariance matrix $\boldsymbol{K}$ given by (48) from Lemma 2 and let $\boldsymbol{\alpha} = (\alpha\mathbf{1}_{(m)}\ \eta\mathbf{1}_{(n)})$ be a vector containing m constant values α, and $n = L - m$ constant values η. Then the probability $\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha})$ can be upper-bounded by

$$\begin{aligned}
\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha}) \leq \int_{\frac{\phi(\alpha,\eta)}{\sqrt{\sigma_L}}}^{\infty} f(y_1) \int_{g(y_1)}^{h(y_1)} f(y_2)\mathrm{d}y_2 \prod_{\ell=1}^{m-1} \left(\int_{v_\ell(y_1)}^{u_\ell(y_1)} f(x)\mathrm{d}x \right)\\
\times \prod_{\ell=m}^{L-2} \left(\int_{w_\ell(y_1)}^{z_\ell(y_1)} f(x)\mathrm{d}x \right) \mathrm{d}y_1. \qquad (59)
\end{aligned}$$

The expressions for the integration limits in the above formula are

$$\begin{aligned}
\phi(\alpha,\eta) &= n(bm\alpha + \eta(\xi_L - \lambda_{\mathrm{A}1}))\\
g(y_1) &= -\frac{y_1\sqrt{\sigma_L}(\lambda_{\mathrm{A}1} - \xi_{L-1}) - bmn\alpha(\xi_L - \xi_{L-1})}{(\xi_L - \lambda_{\mathrm{A}1})\sqrt{\sigma_{L-1}}}\\
h(y_1) &= \frac{y_1\sqrt{\sigma_L} - n\eta(\xi_L - \xi_{L-1})}{\sqrt{\sigma_{L-1}}}
\end{aligned}$$

$$\begin{aligned}
u_\ell(y_1) &= \frac{y_1\sqrt{\sigma_L} - \phi(\alpha,\eta)}{bn\sqrt{\sigma_\ell}}\\
v_\ell(y_1) &= -\frac{u_\ell(y_1)}{\ell}\\
z_\ell(y_1) &= \frac{y_1\sqrt{\sigma_L} - \phi(\alpha,\eta)}{(\xi_L - \lambda_{\mathrm{A}1})\sqrt{\sigma_\ell}}\\
w_\ell(y_1) &= -\frac{z_\ell(y_1)}{\ell - m + 1}
\end{aligned}$$

where the values of σ_ℓ, $\ell = 1,2,\ldots,L$, are defined in the following table

ℓ	σ_ℓ
$1 \leq \ell \leq m-1$	$(a_0 - a)(\ell+1)/\ell$
$m \leq \ell \leq L-2$	$(c_0 - c)(\ell - m + 2)/(\ell - m + 1)$
$\ell = L-1$	$\xi_{L-1}\left(b^2n^2m + n(\xi_{L-1} - \lambda_{\mathrm{A}1})^2\right)$
$\ell = L = m+n$	$\xi_L\left(b^2n^2m + n(\xi_L - \lambda_{\mathrm{A}1})^2\right)$

where $\lambda_{\mathrm{A}1}$, $\lambda_{\mathrm{C}1}$, ξ_{L-1}, and ξ_L are given by (56), (57), (54), and (55), respectively.

Proofs of Theorems 3 and 4 are given in Appendix.

Consider now, for example, the worst-case list configuration for a code with even minimum Hamming distance $d_{\mathrm{H}\min}$. The corresponding matrix $\boldsymbol{\Gamma}$ is specified in Statement 4 of Theorem 2 and the Hamming list distance is

$$d_{\mathrm{HL}\min} = \frac{2L}{L+1} d_{\mathrm{H}\min}.$$

According to (5), the list error probability for the given list is

$$P_{\mathrm{e}L}(\boldsymbol{\Gamma}) = \Pr\left(\boldsymbol{t} \geq 4E_s\frac{d_{\mathrm{H}\min}}{2}\mathbf{1}_{(L)}\right) = \Pr(\boldsymbol{t} \geq 2E_s d_{\mathrm{H}\min}\mathbf{1}_{(L)}).$$

To upper-bound the probability $\Pr(\boldsymbol{t} \geq 2E_s d_{\mathrm{H}\min}\mathbf{1}_{(L)})$ we apply Theorem 3 with $\alpha = 2E_s d_{\mathrm{H}\min}$. The integration limits in the bound (58) from Theorem 3 depend on the eigenvalues of the covariance matrix of $\boldsymbol{t}$, that is, $\boldsymbol{K} = \boldsymbol{\Gamma}N_0/2$. We determine them by applying Lemma 1 with $\beta = 2E_s d_{\mathrm{H}\min}N_0$ and $\kappa = 1/2$. Thus we obtain that the largest eigenvalue of $\boldsymbol{K}$ is $\lambda_1 = E_s d_{\mathrm{H}\min}(L+1)N_0$, while the other $(L-1)$ eigenvalues are equal to $\lambda_\ell = E_s d_{\mathrm{H}\min}N_0$. By substituting these values into (58) we obtain the following bound on the error probability $P_{\mathrm{e}L}(\boldsymbol{\Gamma})$:

$$\begin{aligned}
P_{\mathrm{e}L}(\boldsymbol{\Gamma}) &\leq \int_{\sqrt{2d_{\mathrm{HL}\min}E_s/N_0}}^{\infty} f(y) \prod_{\ell=2}^{L} \left(\int_{v_\ell(y)}^{u_\ell(y)} f(x)\mathrm{d}x \right) \mathrm{d}y\\
&= \int_{\sqrt{2d_{\mathrm{HL}\min}E_s/N_0}}^{\infty} f(y) \prod_{\ell=2}^{L} \left(Q(u_\ell(y)) - Q(v_\ell(y)) \right) \mathrm{d}y
\end{aligned} \qquad (60)$$

where

$$\begin{aligned}
u_\ell &= \sqrt{\frac{L(L+1)(\ell-1)}{\ell}}\left(y - \sqrt{2d_{\mathrm{HL}\min}E_s/N_0}\right)\\
v_\ell &= -\frac{u_\ell}{\ell - 1}, \quad \ell = 2,3,\ldots,L.
\end{aligned}$$

For $L = 2$, bound (60) holds with equality and it is illustrated in Figure 1 as the probability that the received signal is in the

decision region $\mathcal{D}$. In Figure 7 the new bound (60) is plotted for $d_{\mathrm{Hmin}} = 4$ and list sizes $L \in \{2,\, 3,\, 5\}$. For comparison, bound (42) is shown in the same figure.

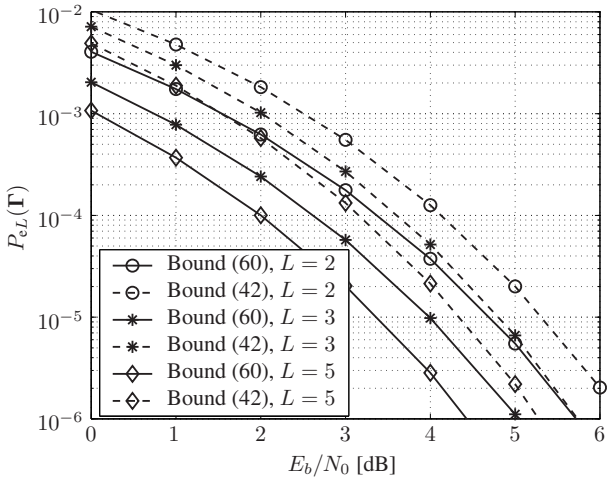

Fig. 7. Comparison of the bounds (60) and (42) on the list error probability for a worst-case list configuration for $d_{\mathrm{Hmin}} = 4$.

IV. Generalized Tangential Bound on the List Error Probability

The bound given by (9) is a generalization of a union bound for list decoding. It upper-bounds the list error probability by a sum of error probabilities $P_{\mathrm{e}L}(\mathbf{\Gamma})$ for all possible list configurations $\mathbf{\Gamma}$ of a given list size. The multiplicities $N(\mathbf{\Gamma})$ of error probabilities $P_{\mathrm{e}L}(\mathbf{\Gamma})$ are, for most practically interesting codes and list sizes, very large numbers (see, *e.g.*, Table III). Hence, the union bound is often not tight even at higher signal-to-noise ratio (SNR) levels.

In order to improve the union-type bound (9) we start from the Fano-Gallager bounding principle [22] which is a commonly used approach for constructing good bounds on the error probability of ML decoding (*cf.* [23], [24]). Namely, if ε denotes a list error event, then the probability of list error $\Pr(\varepsilon)$ can be decomposed and upper-bounded as

$$\begin{aligned} P_{\mathrm{e}L} \equiv \Pr(\varepsilon) &= \Pr\left(\varepsilon, \boldsymbol{n} \in \mathcal{A}\right) + \Pr\left(\varepsilon, \boldsymbol{n} \notin \mathcal{A}\right) \\ &= \Pr\left(\varepsilon, \boldsymbol{n} \in \mathcal{A}\right) + \Pr\left(\varepsilon | \boldsymbol{n} \notin \mathcal{A}\right) \Pr(\boldsymbol{n} \notin \mathcal{A}) \\ &\leq \Pr\left(\varepsilon, \boldsymbol{n} \in \mathcal{A}\right) + \Pr(\boldsymbol{n} \notin \mathcal{A}) \end{aligned} \tag{61}$$

where $\mathcal{A}$ denotes a subset of noise vectors $\boldsymbol{n}$. The region $\mathcal{A}$ and its complement $\mathcal{A}^{\mathrm{c}}$ are interpreted as the regions of few and many errors, respectively. In the above expansion, the probability $\Pr(\varepsilon | \boldsymbol{n} \in \mathcal{A}^{\mathrm{c}})$ is upper-bounded by 1, which yields the bound (61). Further bounding of (61) is obtained by upper-bounding the error probability inside the region $\mathcal{A}$, $\Pr\left(\varepsilon, \boldsymbol{n} \in \mathcal{A}\right)$. Clearly, the choice of the region $\mathcal{A}$ influences the tightness of the bound.

We adopt the tangential-bound approach from [17] and decompose the noise vector $\boldsymbol{n}$ into the radial component (along the transmitted signal $\boldsymbol{s}_0$) and L components orthogonal to the radial component. Then let $\mathcal{A}$ be a set of noise vectors $\boldsymbol{n}$ for which the radial component is smaller than or equal to a given threshold T. In this case inequality (61) can be rewritten as

$$P_{\mathrm{e}L} \leq \Pr\left(\varepsilon, \langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle \leq T\right) + \Pr\left(\langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle > T\right). \tag{62}$$

Since the radial noise component is a Gaussian random variable, the probability $\Pr\left(\langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle > T\right)$ is simply a Q-function of the threshold T. The exact computation of the first term of (62) is infeasible; in the following we will develop an upper bound for it.

The list error condition (3) is equivalent to

$$\langle \boldsymbol{n}, \boldsymbol{s}_\ell \rangle \geq \langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle + d_{\mathrm{E}}^2(\boldsymbol{s}_0, \boldsymbol{s}_\ell)/2. \tag{63}$$

For a given value x of the radial noise component, by introducing notation $z_\ell = \langle \boldsymbol{n}, \boldsymbol{s}_\ell \rangle$, $\ell = 0, 1, \ldots, L$, the list error condition (63) can be rewritten as

$$\begin{cases} z_0 = \langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle = x \\ z_\ell = \langle \boldsymbol{n}, \boldsymbol{s}_\ell \rangle \geq x + 2E_s d_{\mathrm{H}0\ell}, \quad \ell = 1, 2, \ldots, L \end{cases} \tag{64}$$

where $d_{\mathrm{H}0\ell} = d_{\mathrm{H}}(\boldsymbol{v}_0, \boldsymbol{v}_\ell)$. The vector $\boldsymbol{z} = (z_0\, z_1\, \ldots\, z_L)$ has Gaussian distribution with zero mean and covariance matrix

$$\boldsymbol{K}_z = \frac{N_0}{2} N E_s \{k_{ij}\}, \quad k_{ij} = \begin{cases} 1, & i = j \\ 1 - 2\delta_{ij}, & i \neq j \end{cases} \tag{65}$$

where $i, j = 0, 1, \ldots, L$ and $\delta_{ij} = \delta_{ji} = d_{\mathrm{H}ij}/N$ is the relative Hamming distance between the ith and jth codewords.

Next, our goal is to orthogonalize the radial noise component z_0 with respect to the other L components z_ℓ. Similarly as in [28], we introduce transformed variables $\boldsymbol{y} = (y_1\, y_2\, \ldots\, y_L)$ such that they are uncorrelated with z_0. The components y_ℓ are given by

$$y_\ell = z_\ell - (1 - 2\delta_{0\ell}) z_0, \quad \ell = 1, 2, \ldots, L. \tag{66}$$

Then, for a given $z_0 = x$, the error condition (64) is equivalent to

$$y_\ell \geq 2\delta_{0\ell}(x + NE_s), \quad \ell = 1, 2, \ldots, L. \tag{67}$$

Indeed, the components y_ℓ are uncorrelated with the component z_0, that is,

$$\begin{aligned} \mathrm{E}\left[z_0 y_\ell\right] &= \mathrm{E}\left[z_0 z_\ell\right] - (1 - 2\delta_{0\ell})\, \mathrm{E}\left[z_0^2\right] \\ &= (1 - 2\delta_{0\ell}) N E_s N_0/2 - (1 - 2\delta_{0\ell}) N E_s N_0/2 \\ &= 0. \end{aligned}$$

The covariance matrix of the vector $\boldsymbol{y}$ is equal to

$$\boldsymbol{K}_y = 2N_0 N E_s \{k_{ij}\}, \quad k_{ij} = \frac{\delta_{0i} + \delta_{0j} - \delta_{ij}}{2} - \delta_{0i}\delta_{0j} \tag{68}$$

where $i, j = 1, 2, \ldots, L$. Then, according to (62) the list error probability $P_{\mathrm{e}L}$ can be upper-bounded by

$$P_{\mathrm{e}L} \leq \min_T \left\{ \int_{-\infty}^{T} f(x) \min\left\{ 1, \sum_{\boldsymbol{K}_y} N(\boldsymbol{K}_y) P_{\mathrm{e}L}(\boldsymbol{K}_y, x) \right\} \mathrm{d}x + Q(T) \right\} \tag{69}$$

where we have used a union-type bound for the first term in (62). The minimization over threshold values T is performed in order to obtain the tightest upper bound. $N(\boldsymbol{K}_y)$ is the

number of lists that yield the covariance matrix $\boldsymbol{K}_y$ and $P_{\mathrm{e}L}(\boldsymbol{K}_y, x)$ is the error probability for the list described by the covariance matrix $\boldsymbol{K}_y$, under the condition that the radial noise level is $z_0 = x$. From (67) it follows that

$$P_{\mathrm{e}L}(\boldsymbol{K}_y, x) = \Pr(\boldsymbol{y} \geq 2(x + NE_s)\boldsymbol{\delta}_0) \tag{70}$$

where $\boldsymbol{\delta}_0 = (\delta_{01}\, \delta_{02}\, \ldots\, \delta_{0L})$. For the worst-case list configurations, with even and odd minimum distance, the error probability (70) can be upper-bounded using Theorems 3 and 4, respectively.

We now consider linear block codes with even minimum Hamming distance. For rather high signal-to-noise ratios the dominant term in (69) corresponds to the worst-case list configuration, where all codewords have relative Hamming weight $\delta = d_{\mathrm{H}\min}/N$, and they are at the relative pairwise distances $\delta_{ij} = \delta$, $i \neq j$. In this case, the covariance matrix $\boldsymbol{K}_y$ from (68) simplifies to

$$\boldsymbol{K}_y = 2N_0NE_s\{k_{ij}\}, \quad k_{ij} = \begin{cases} \delta(1-\delta), & i = j \\ (1-2\delta)/2, & i \neq j \end{cases}$$

for $i, j = 1, 2, \ldots, L$. The eigenvalues of $\boldsymbol{K}_y$ are obtained by applying Lemma 1 with $\beta = 2N_0NE_s\delta(1-\delta)$ and $\kappa = \frac{1-2\delta}{2(1-\delta)}$. The dominant error term (70)

$$P_{\mathrm{e}L}(\boldsymbol{K}_y, x) = \Pr(\boldsymbol{y} \geq 2(x + NE_s)\delta\mathbf{1}_{(L)})$$

can be upper-bounded by applying Theorem 3 with $\alpha = 2(x + NE_s)\delta$. For example, for a list of size $L = 2$ we obtain

$$P_{\mathrm{e}L}(\boldsymbol{K}_y, x) = \int\limits_{\frac{2\sqrt{2\delta}}{\sqrt{N_0NE_s(3-4\delta)}}(x+NE_s)}^{\infty} f(y)\left(Q(u(x,y)) - Q(-u(x,y))\right)\mathrm{d}y$$

where

$$u(x,y) = y\sqrt{(3-4\delta)} - \frac{2\sqrt{2\delta}}{\sqrt{N_0NE_s}}(x + NE_s).$$

As follows from the examples shown in Tables I–III, list configuration matrices corresponding to other terms of the bound (69) can have rather complicated structures. In order to simplify the computations, for the rest of the terms we use bounds which, like the bound (42), are based only on the list distances $d_{\mathrm{H}L}(\boldsymbol{\Gamma})$ and do not take into account the fine structure of the Gram matrix $\boldsymbol{\Gamma}$.

In the case of list size $L = 1$, the bound (69) reduces to

$$P_{\mathrm{e}1} \leq \min_T \left\{ \int\limits_{-\infty}^{T} f(x) \min\left\{1, \sum_{d_{\mathrm{H}}} N(d_{\mathrm{H}}) P_{\mathrm{e}}(d_{\mathrm{H}}, x)\right\} \mathrm{d}x + Q(T) \right\} \tag{71}$$

where $N(d_{\mathrm{H}})$ are the spectrum coefficients of the code and $P_{\mathrm{e}}(d_{\mathrm{H}}, x)$, according to Theorem 3, equals

$$P_{\mathrm{e}}(d_{\mathrm{H}}, x) = Q\left(\sqrt{\frac{2d_{\mathrm{H}}}{N_0NE_s(N - d_{\mathrm{H}})}}\,(x + NE_s)\right). \tag{72}$$

As an analogy to (71) we have the tangential bound for list decoding

$$P_{\mathrm{e}L} \leq \min_T \left\{ \int\limits_{-\infty}^{T} f(x) \min\left\{1, \sum_{\boldsymbol{\Gamma}} N(\boldsymbol{\Gamma}) P_{\mathrm{e}}(d_{\mathrm{H}L}(\boldsymbol{\Gamma}), x)\right\} \mathrm{d}x + Q(T) \right\} \tag{73}$$

where the terms $P_{\mathrm{e}}(d_{\mathrm{H}L}(\boldsymbol{\Gamma}), x)$ are given by (72). Finally, by combining bounds (69) and (73) we obtain our new generalized tangential bound

$$P_{\mathrm{e}L} \leq \min_T \left\{ \int\limits_{-\infty}^{T} f(x) \min\left\{1, N(\boldsymbol{K}_y) P_{\mathrm{e}L}(\boldsymbol{K}_y, x) + \sum_{\boldsymbol{\Gamma}} N(\boldsymbol{\Gamma}) P_{\mathrm{e}L}(d_{\mathrm{H}L}(\boldsymbol{\Gamma}), x)\right\} \mathrm{d}x + Q(T) \right\} \tag{74}$$

where the dominant term is estimated as in (70), while the remaining terms are upper-bounded only by using the list distance $d_{\mathrm{H}L}(\boldsymbol{\Gamma})$, which is equivalent to replacing the codeword sets with list configuration $\boldsymbol{\Gamma}$ by an "average codeword" at distance $d_{\mathrm{H}L}(\boldsymbol{\Gamma})$ from the transmitted codeword.

We illustrate the obtained bounds using the $(24, 12, 8)$ Golay code and list decoding with the list size $L = 2$. The list configuration matrices and their multiplicities are shown in Table III. In Figure 8 the union bound (9), the tangential bound (73), and the improved tangential bound (74) are shown and compared with the list error probability obtained by simulations. When computing the union bound (9), each error probability for a given list, $P_{\mathrm{e}L}(\boldsymbol{\Gamma})$, is upper-bounded by a Q-function given by (42). Note that bound (42) is looser than bound (60) for the worst-case list configuration (which dominates the performance for high SNR), as already illustrated in Figure 7. Moreover, from Table III it follows that the multiplicities $N(\boldsymbol{\Gamma})$ of the possible list configuration matrices are very large—there are more than 10^5 codeword pairs with the worst-case list configuration yielding minimum list distance. Therefore, union bound (9) is quite loose, even at higher SNR. The new bound (74) is significantly tighter than the union bound in the whole range of the observed SNR levels.

Further improvement of the new bound, for example, using the TSB technique, is a challenging problem which is a topic of future research.

V. Conclusions

In this paper, list decoding is considered. The so-called minimum list distance of a signal constellation is defined, which plays the same role for list decoding as the minimum distance for maximum-likelihood decoding. It was shown that the list distance is determined by the list configuration matrix, which is a Gram matrix obtained from the signal vectors forming the constellation. Several properties of the list configuration matrix are established. Using these properties,

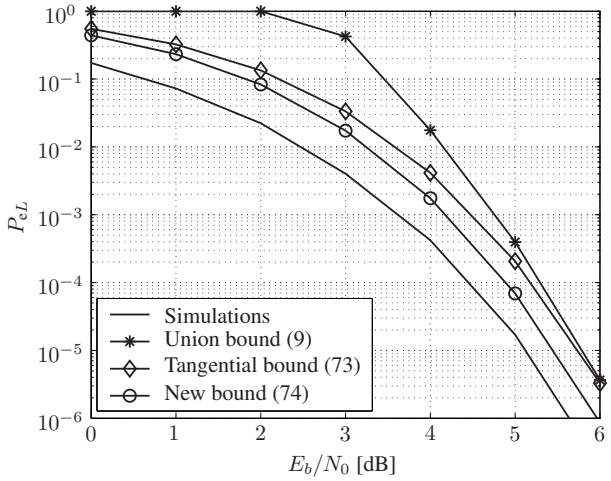

Fig. 8. Comparison of bounds (9), (73), and (74) for list decoding of the $(24, 12, 8)$ Golay code with list size $L = 2$.

the minimum list distance is computed for binary linear codes with even and odd minimum distance. For both cases, the worst-case list configurations, which yield the minimum list distances are identified and a new upper bound on the list error probability for a given worst-case list is derived. Furthermore, an improved tangential union bound is derived and by combining it with the bound for a given worst-case list a new upper bound on the list decoding error probability is obtained. The new bound is tighter than previously known union bound.

VI. Acknowledgement

The authors are grateful to both reviewers for their constructive comments which improved the presentation of these results. In particular, they wish to thank the reviewer who pointed out the observations presented in Section II.D and who brought to their attention references [12], [16], and [30]. The authors also wish to thank Thomas Magesacher (Lund University) for his help in the preparation of this manuscript.

VII. Appendix

A. Proof of Lemma 1

The eigenvalues λ_ℓ, $\ell = 1, 2, \ldots, L$, of the matrix $\boldsymbol{K}$ are the solutions of the characteristic equation

$$\det(\boldsymbol{K} - \lambda \boldsymbol{I}) = \begin{vmatrix} \beta - \lambda & \beta\kappa & \ldots & \beta\kappa \\ \beta\kappa & \beta - \lambda & \ldots & \beta\kappa \\ \vdots & \vdots & \ddots & \vdots \\ \beta\kappa & \beta\kappa & \ldots & \beta - \lambda \end{vmatrix} = 0.$$

The above determinant is easily obtained by, for example, transforming the matrix into triangular form, whose determinant is then the product of the diagonal elements. Thus we obtain

$$(\beta(1 + \kappa(L-1)) - \lambda) \cdot (\beta(1-\kappa) - \lambda)^{L-1} = 0$$

which yields the solutions λ_1 as in (44) and λ_ℓ, $\ell = 2, 3, \ldots, L$, as in (45).

The eigenvectors $\boldsymbol{x}_\ell = (x_\ell^{(1)}\, x_\ell^{(2)}\, \ldots\, x_\ell^{(L)})$ corresponding to the eigenvalues λ_ℓ, $\ell = 1, 2, \ldots, L$, of the matrix $\boldsymbol{K}$ satisfy the equation

$$\lambda_\ell \boldsymbol{x}_\ell = \boldsymbol{x}_\ell \boldsymbol{K}. \tag{75}$$

For $\lambda_1 = \beta(1 + \kappa(L-1))$, (75) yields

$$x_1^{(1)} = x_1^{(2)} = \cdots = x_1^{(L)} = \frac{1}{L} \sum_{i=1}^{L} x_1^{(i)}.$$

Hence, any vector with L identical components can be chosen as an eigenvector. Without loss of generality we choose $\boldsymbol{x}_1 = (1\, 1\, \ldots\, 1) = \mathbf{1}_{(L)}$.

For $\lambda_\ell = \beta(1 - \kappa)$, $\ell = 2, 3, \ldots, L$, (75) yields

$$\sum_{i=1}^{L} x_\ell^{(i)} = 0.$$

Hence, any set of $(L-1)$ linearly independent vectors, such that their elements sum up to 0, is a set of eigenvectors. We choose

$$\begin{aligned} \boldsymbol{x}_\ell &= \Big(\underbrace{-\frac{1}{\ell-1} \; -\frac{1}{\ell-1} \; \ldots \; -\frac{1}{\ell-1}}_{\ell-1 \text{ times}} \; 1\; 0 \ldots 0 \Big) \\ &= \frac{1}{\ell-1} (-\mathbf{1}_{(\ell-1)}\; \ell-1\; \mathbf{0}_{(L-\ell)}), \quad \ell = 2, 3, \ldots, L \end{aligned}$$

which completes the proof. ■

Remark 1: The equation (75) can be rewritten in a matrix form as

$$\boldsymbol{K} = \boldsymbol{V} \boldsymbol{\Lambda} \boldsymbol{V}^{-1}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1\, \lambda_2\, \ldots\, \lambda_L)$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{K}$, and $\boldsymbol{V}$ is the eigenvector matrix whose columns are eigenvectors $\boldsymbol{x}_\ell^{\mathrm{T}}$, $\ell = 1, 2, \ldots L$, that is,

$$\begin{aligned} \boldsymbol{V} &= \left(\boldsymbol{x}_1^{\mathrm{T}}\, \boldsymbol{x}_2^{\mathrm{T}}\, \ldots\, \boldsymbol{x}_L^{\mathrm{T}}\right) \\ &= \begin{pmatrix} 1 & -1 & -1/2 & -1/3 & \ldots & -1/(L-1) \\ 1 & 1 & -1/2 & -1/3 & \ldots & -1/(L-1) \\ 1 & 0 & 1 & -1/3 & \ldots & -1/(L-1) \\ 1 & 0 & 0 & 1 & \ldots & -1/(L-1) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}. \end{aligned} \tag{76}$$

Since $\boldsymbol{K}$ is a normal matrix, its eigenvector matrix $\boldsymbol{V}$ describes an orthogonal transform, that is,

$$\boldsymbol{V}^{\mathrm{T}} \boldsymbol{V} = \boldsymbol{D}$$

where $\boldsymbol{D} = \mathrm{diag}(d_1\; d_2\; \ldots\; d_L)$ is the diagonal matrix with elements

$$d_\ell = \begin{cases} L, & \ell = 1 \\ \ell/(\ell-1), & \ell = 2, 3, \ldots, L. \end{cases}$$

This orthogonal transform reduces the matrix $\boldsymbol{K}$ to the diagonal matrix:

$$\boldsymbol{V}^{\mathrm{T}} \boldsymbol{K} \boldsymbol{V} = \boldsymbol{D}\boldsymbol{\Lambda} = \mathrm{diag}(\sigma_1\; \sigma_2\; \ldots\; \sigma_L) \tag{77}$$

where

$$\sigma_\ell = \lambda_\ell d_\ell = \begin{cases} L\lambda_1, & \ell = 1 \\ \lambda_\ell \ell/(\ell-1), & \ell = 2,3,\ldots,L. \end{cases} \quad (78)$$

This transform will be used for proving Theorem 3.

B. Proof of Lemma 2

The matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ from (49) and (50) have the same general structure as the matrix (43) from Lemma 1 with $\beta = a_0$, $\kappa = a/a_0$, and $\beta = c_0$, $\kappa = c/c_0$, respectively. Hence, it follows from (44)-(45) that the eigenvalues of $\boldsymbol{A}$ are

$$\begin{aligned} \lambda_{\mathrm{A}1} &= a_0 + a(m-1) \\ \lambda_{\mathrm{A}i} &= a_0 - a, \quad i = 2,3,\ldots,m \end{aligned}$$

and the corresponding eigenvectors are

$$\boldsymbol{x}_{\mathrm{A}1} = \mathbf{1}_{(m)} \quad (79)$$

$$\boldsymbol{x}_{\mathrm{A}i} = \frac{1}{i-1}\left(-\mathbf{1}_{(i-1)} \;\; i-1 \;\; \mathbf{0}_{(m-i)}\right) \quad (80)$$

where $i = 2,3,\ldots,m$. Similarly, the eigenvalues of $\boldsymbol{C}$ are

$$\begin{aligned} \lambda_{\mathrm{C}1} &= c_0 + c(n-1) \\ \lambda_{\mathrm{C}j} &= c_0 - c \end{aligned}$$

for $j = 2,3,\ldots,n$, and the corresponding eigenvectors are

$$\boldsymbol{x}_{\mathrm{C}1} = \mathbf{1}_{(n)} \quad (81)$$

$$\boldsymbol{x}_{\mathrm{C}j} = \frac{1}{j-1}\left(-\mathbf{1}_{(j-1)} \;\; j-1 \;\; \mathbf{0}_{(n-j)}\right) \quad (82)$$

where $j = 2,3,\ldots,n$ and $n = L-m$. Now consider the $L \times L$ block-diagonal matrix

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C} \end{pmatrix}.$$

Its eigenvalues are simply those of $\boldsymbol{A}$ and $\boldsymbol{C}$, and its eigenvectors are

$$\begin{aligned} \boldsymbol{x}_{\mathrm{G}i} &= (\boldsymbol{x}_{\mathrm{A}i} \; \mathbf{0}_{(n)}), \quad i = 1,2,\ldots,m \\ \boldsymbol{x}_{\mathrm{G}m+j} &= (\mathbf{0}_{(m)} \; \boldsymbol{x}_{\mathrm{C}j}), \quad j = 1,2,\ldots,n. \end{aligned}$$

All the vectors $\boldsymbol{x}_{\mathrm{A}i}$ except the all-one vector, $\boldsymbol{x}_{\mathrm{A}1} = \mathbf{1}_{(m)}$, have the property that their components sum up to zero. Hence, for those $m-1$ vectors we have

$$\begin{aligned} (\boldsymbol{x}_{\mathrm{A}i} \; \mathbf{0}_{(n)})\boldsymbol{K} &= (\boldsymbol{x}_{\mathrm{A}i} \; \mathbf{0}_{(n)}) \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{C} \end{pmatrix} \\ &= (\boldsymbol{x}_{\mathrm{A}i}\boldsymbol{A} \;\; \boldsymbol{x}_{\mathrm{A}i}\boldsymbol{B}) = \lambda_{\mathrm{A}i}(\boldsymbol{x}_{\mathrm{A}i} \; \mathbf{0}_{(n)}) \end{aligned}$$

which follows from the fact that all elements of the matrix $\boldsymbol{B}$ are the same. We conclude that the eigenvectors $(\boldsymbol{x}_{\mathrm{A}i} \; \mathbf{0}_{(n)})$, $i = 2,3,\ldots,m$, are also eigenvectors of $\boldsymbol{K}$ and the corresponding eigenvalues are the eigenvalues of $\boldsymbol{A}$, that is, $\lambda_{\mathrm{A}i} = a_0 - a$.

Equivalently, we obtain that the vectors $(\mathbf{0}_{(m)} \; \boldsymbol{x}_{\mathrm{C}j})$, except the one with $\boldsymbol{x}_{\mathrm{C}1} = \mathbf{1}_{(n)}$, are also eigenvectors of $\boldsymbol{K}$ and the corresponding eigenvalues are the eigenvalues of $\boldsymbol{C}$, that is, $\lambda_{\mathrm{C}j} = c_0 - c$, $j = 2,3,\ldots,n$.

Thus, we have determined $m+n-2 = L-2$ eigenvalues of $\boldsymbol{K}$:

$$\begin{aligned} \xi_\ell &= a_0 - a, \quad \ell = 1,2,\ldots,m-1 \\ \xi_\ell &= c_0 - c, \quad \ell = m, m+1,\ldots,L-2 \end{aligned}$$

and the corresponding eigenvectors

$$\boldsymbol{x}_\ell = (\boldsymbol{x}_{\mathrm{A}\ell+1} \, \mathbf{0}_{(n)}), \; \ell = 1,2,\ldots,m-1 \quad (83)$$

$$\boldsymbol{x}_\ell = (\mathbf{0}_{(m)} \, \boldsymbol{x}_{\mathrm{C}\ell-m+2}), \; \ell = m, m+1,\ldots,L-2. \quad (84)$$

where $\boldsymbol{x}_{\mathrm{A}i}$ and $\boldsymbol{x}_{\mathrm{C}j}$ are given by (80) and (82), respectively.

Now we only need to determine the remaining two eigenvalues ξ_{L-1} and ξ_L, and the corresponding eigenvectors. To this end, we first exploit the property that the trace of any square matrix is equal to the sum of its eigenvalues. Thus we have

$$\mathrm{tr}(\boldsymbol{G}) = \sum_{i=1}^{m} \lambda_{\mathrm{A}i} + \sum_{j=1}^{n} \lambda_{\mathrm{C}j}$$

and also

$$\mathrm{tr}(\boldsymbol{K}) = \sum_{i=2}^{m} \lambda_{\mathrm{A}i} + \sum_{j=2}^{n} \lambda_{\mathrm{C}j} + \xi_{L-1} + \xi_L.$$

Since $\mathrm{tr}(\boldsymbol{G}) = \mathrm{tr}(\boldsymbol{K})$, we obtain that the sum of two unknown eigenvalues is

$$\xi_{L-1} + \xi_L = \lambda_{\mathrm{A}1} + \lambda_{\mathrm{C}1} = a_0 + c_0 + (m-1)a + (n-1)c. \quad (85)$$

The determinant of the matrix $\boldsymbol{K}$ can be written as (*cf.* [31])

$$\det(\boldsymbol{K}) = \det(\boldsymbol{A}) \det(\boldsymbol{C} - \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}). \quad (86)$$

The inverse of the matrix $\boldsymbol{A}$ can be written as

$$\boldsymbol{A}^{-1} = \boldsymbol{V}_{\mathrm{A}} \boldsymbol{\Lambda}_{\mathrm{A}}^{-1} \boldsymbol{V}_{\mathrm{A}}^{-1}$$

where $\boldsymbol{\Lambda}_{\mathrm{A}}^{-1} = \mathrm{diag}\,(1/\lambda_{\mathrm{A}1} \; 1/\lambda_{\mathrm{A}2} \; \ldots \; 1/\lambda_{\mathrm{A}m})$ and $\boldsymbol{V}_{\mathrm{A}}$ is the eigenvector matrix of the form (76), whose columns are eigenvectors $\boldsymbol{x}_{\mathrm{A}i}^{\mathrm{T}}$, $i = 1,2,\ldots,m$. Since the all-one vector $\boldsymbol{x}_{\mathrm{A}1}^{\mathrm{T}}$ is the first column of $\boldsymbol{V}_{\mathrm{A}}$, and $\det(\boldsymbol{V}_{\mathrm{A}}) = 1$, we easily obtain that

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B} = \frac{b^2m^2}{\lambda_{\mathrm{A}1}} \mathbf{1}_{(n\times n)}$$

where $\mathbf{1}_{(n\times n)}$ denotes the all-one matrix of size $n \times n$. Thus, the eigenvalues of the matrix $\boldsymbol{C} - \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}$ can be obtained from Lemma 1 with $\beta = c_0 - b^2m^2/\lambda_{\mathrm{A}1}$ and $\beta\kappa = c - b^2m^2/\lambda_{\mathrm{A}1}$, which yields the largest eigenvalue equal to $\lambda_{\mathrm{C}1} - b^2m^2n/\lambda_{\mathrm{A}1}$, while the remaining eigenvalues are equal to $\lambda_{\mathrm{C}j} = c_0 - c$, $j = 2,3,\ldots,n$.

Since the determinant of the matrix is equal to the product of its eigenvalues, (86) is equivalent to

$$\xi_{L-1}\xi_L \prod_{i=2}^{m} \lambda_{\mathrm{A}i} \prod_{j=2}^{n} \lambda_{\mathrm{C}j} = \prod_{i=1}^{m} \lambda_{\mathrm{A}i} \prod_{j=2}^{n} \lambda_{\mathrm{C}j} \left(\lambda_{\mathrm{C}1} - \frac{b^2m^2n}{\lambda_{\mathrm{A}1}}\right)$$

which yields

$$\xi_{L-1}\xi_L = \lambda_{\mathrm{A}1}\left(\lambda_{\mathrm{C}1} - \frac{b^2m^2n}{\lambda_{\mathrm{A}1}}\right). \quad (87)$$

From (85) and (87), which specify the sum and the product of two last eigenvalues of $\boldsymbol{K}$, we obtain

$$\begin{aligned}\xi_{L-1} &= \frac{\lambda_{\mathrm{A}1}+\lambda_{\mathrm{C}1}}{2}-\frac{1}{2}\sqrt{(\lambda_{\mathrm{A}1}-\lambda_{\mathrm{C}1})^2+4b^2m^2n}\\ \xi_{L} &= \frac{\lambda_{\mathrm{A}1}+\lambda_{\mathrm{C}1}}{2}+\frac{1}{2}\sqrt{(\lambda_{\mathrm{A}1}-\lambda_{\mathrm{C}1})^2+4b^2m^2n}.\end{aligned}$$

The eigenvectors corresponding to ξ_{L-1}, ξ_L satisfy the equation

$$\xi_\ell \boldsymbol{K} = \boldsymbol{x}_\ell \boldsymbol{K}, \quad \ell = L-1,\, L$$

which imposes the following structure of the eigenvectors:

$$\boldsymbol{x}_\ell = \left(\varrho_1 \mathbf{1}_{(m)} \;\; \varrho_2 \mathbf{1}_{(n)}\right), \quad \varrho_2 = \varrho_1 \frac{\xi_\ell - \lambda_{\mathrm{A}1}}{bn}, \quad \ell = L-1,\, L$$

for any arbitrary ϱ_1. Without loss of generality, we can choose $\varrho_1 = bn$ and thus we obtain

$$\boldsymbol{x}_{L-1} = \left(bn\cdot\mathbf{1}_{(m)} \;\; (\xi_{L-1}-\lambda_{\mathrm{A}1})\cdot\mathbf{1}_{(n)}\right) \tag{88}$$

$$\boldsymbol{x}_{L} = \left(bn\cdot\mathbf{1}_{(m)} \;\; (\xi_{L}-\lambda_{\mathrm{A}1})\cdot\mathbf{1}_{(n)}\right) \tag{89}$$

which concludes the proof. ■

Remark 2: The matrix $\boldsymbol{K}$ can be written in the form

$$\boldsymbol{K} = \boldsymbol{T}\boldsymbol{\Psi}\boldsymbol{T}^{-1}$$

where $\boldsymbol{\Psi} = \mathrm{diag}\,(\xi_1\, \xi_2\, \ldots\, \xi_L)$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{K}$, and $\boldsymbol{T}$ is the eigenvector matrix whose columns are the eigenvectors $\boldsymbol{x}_\ell^{\mathrm{T}}$, $\ell = 1, 2, \ldots, L$. From (83), (84), (88), and (89), it follows that the structure of the matrix $\boldsymbol{T}$ is

$$\boldsymbol{T} = \left(\begin{array}{c|c|c|c} \boldsymbol{X}_{\mathrm{A}} & \mathbf{0}_{(m\times(n-1))} & \boldsymbol{x}_{L-1}^{\mathrm{T}} & \boldsymbol{x}_{L}^{\mathrm{T}} \\ \mathbf{0}_{(n\times(m-1))} & \boldsymbol{X}_{\mathrm{C}} & & \end{array}\right) \tag{90}$$

where $\boldsymbol{X}_{\mathrm{A}}$ is the $m\times(m-1)$ matrix whose columns are $m-1$ eigenvectors of the matrix $\boldsymbol{A}$, $\boldsymbol{x}_{\mathrm{A}i}^{\mathrm{T}}$, $i = 2, \ldots, m$, and $\boldsymbol{X}_{\mathrm{C}}$ is the $n \times (n-1)$ matrix whose columns are the eigenvectors $\boldsymbol{x}_{\mathrm{C}j}^{\mathrm{T}}$, $j = 2, \ldots, n$. The matrix $\boldsymbol{T}$ describes an orthogonal transform, that is

$$\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T} = \boldsymbol{D}$$

where $\boldsymbol{D} = \mathrm{diag}(d_1\, d_2\, \ldots\, d_L)$ is the diagonal matrix with elements

$$d_\ell = \begin{cases} (\ell+1)/\ell, & \ell = 1, 2, \ldots, m-1 \\ (\ell-m+2)/(\ell-m+1), & \ell = m,\, m+1, \ldots, L-2 \\ b^2n^2m + n(\xi_{L-1}-\lambda_{\mathrm{A}1})^2, & \ell = L-1 \\ b^2n^2m + n(\xi_{L}-\lambda_{\mathrm{A}1})^2, & \ell = L. \end{cases}$$

This orthogonal transform reduces matrix $\boldsymbol{K}$ to the diagonal matrix:

$$\boldsymbol{T}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{T} = \boldsymbol{D}\boldsymbol{\Psi} = \mathrm{diag}(\sigma_1\; \sigma_2\; \ldots\; \sigma_L) \tag{91}$$

where

$$\sigma_\ell = \xi_\ell d_\ell = \begin{cases} (a_0-a)(\ell+1)/\ell, \;\; \ell = 1, 2, \ldots, m-1 \\ (c_0-c)(\ell-m+2)/(\ell-m+1), \\ \qquad\qquad \ell = m,\, m+1, \ldots, L-2 \\ \xi_{L-1}\left(b^2n^2m + n(\xi_{L-1}-\lambda_{\mathrm{A}1})^2\right), \;\; \ell = L-1 \\ \xi_{L}\left(b^2n^2m + n(\xi_{L}-\lambda_{\mathrm{A}1})^2\right), \;\; \ell = L. \end{cases} \tag{92}$$

This transform will be used for proving Theorem 4.

C. Proof of Theorem 3

Direct estimation of the error probability $\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha})$ would involve integration of the L-dimensional density function $f(\boldsymbol{t})$ over the region where all $t_\ell \geq \alpha$, $\ell = 1, 2, \ldots, L$. Consider instead the vector $\boldsymbol{q}$ obtained by the orthogonal transformation

$$\boldsymbol{q} = \boldsymbol{t}\boldsymbol{V} \tag{93}$$

where $\boldsymbol{V}$ is the eigenvector matrix of $\boldsymbol{K}$, given by (76). The covariance matrix of $\boldsymbol{q}$ is given by (77). Hence, it follows that q_ℓ, $\ell = 1, 2, \ldots, L$, are independent Gaussian random variables with variances σ_ℓ given by (78). Therefore, the L-dimensional density function $f(\boldsymbol{q})$ decomposes into a product of L marginal densities $f(q_\ell)$, $\ell = 1, 2, \ldots, L$, which simplifies the estimation of the error probability. Now we need to find the integration limits for each q_ℓ, $\ell = 1, 2, \ldots, L$. To this end, we rewrite the matrix equation (93) as

$$q_1 = \sum_{i=1}^{L} t_i \tag{94}$$

$$q_\ell = t_\ell - \frac{1}{\ell-1}\sum_{i=1}^{\ell-1} t_i, \quad \ell = 2, 3, \ldots, L. \tag{95}$$

From $\boldsymbol{t} \geq \boldsymbol{\alpha}$ and (94) it immediately follows that

$$q_1 \geq L\alpha. \tag{96}$$

For the remaining $L-1$ components q_ℓ, $\ell = 2, 3, \ldots, L$, consider the linear combinations

$$\begin{aligned} q_1 + \varphi q_\ell &= \sum_{i=1}^{L} t_i + \varphi t_\ell - \frac{\varphi}{\ell-1}\sum_{i=1}^{\ell-1} t_i \\ &= \left(1-\frac{\varphi}{\ell-1}\right)\sum_{i=1}^{\ell-1} t_i + (1+\varphi)t_\ell + \sum_{i=\ell+1}^{L} t_i. \end{aligned} \tag{97}$$

For $-1 \leq \varphi \leq \ell-1$, all coefficients multiplying t_ℓ, $\ell = 1, 2, \ldots, L$, in (97) are positive and their sum is

$$\left(1-\frac{\varphi}{\ell-1}\right)(\ell-1) + (1+\varphi) + (L-\ell) = L.$$

Therefore, since $t_\ell \geq \alpha$, we have

$$q_1 + \varphi q_\ell \geq \alpha L, \quad \ell = 2, 3, \ldots, L$$

which yields

$$\begin{cases} q_\ell \leq \frac{q_1-\alpha L}{|\varphi|}, & -1 \leq \varphi < 0 \\ q_\ell \geq -\frac{q_1-\alpha L}{\varphi}, & 0 < \varphi \leq \ell-1. \end{cases}$$

The values $\varphi = -1$ and $\varphi = \ell-1$, respectively, minimize the domain of integration. Thus, we obtain that

$$-\frac{q_1-\alpha L}{\ell-1} \leq q_\ell \leq q_1 - \alpha L. \tag{98}$$

Finally, in order to have the Gaussian $\mathcal{N}(0,1)$ distribution, we change the variables according to $y = q_1/\sqrt{(\sigma_1)}$ and $x_\ell = q_\ell/\sqrt{(\sigma_\ell)}$, $\ell = 2, 3, \ldots, L$, and obtain from (96) and (98):

$$\begin{aligned} \frac{\alpha L}{\sqrt{\sigma_1}} &\leq y \leq \;\; \infty \\ -\frac{y\sqrt{\sigma_1}-\alpha L}{(\ell-1)\sqrt{\sigma_\ell}} &\leq x_\ell \leq \;\; \frac{y\sqrt{\sigma_1}-\alpha L}{\sqrt{\sigma_\ell}}, \qquad \ell = 2, 3, \ldots, L \end{aligned}$$

which yields the bound (58) in Theorem 3. ■

D. Proof of Theorem 4

Following the same approach as in the proof of Theorem 3, we consider the vector

$$\boldsymbol{q} = \boldsymbol{t}\boldsymbol{T} \tag{99}$$

where $\boldsymbol{T}$ is given by (90). From (91) it follows that q_ℓ, $\ell = 1, 2, \ldots, L$, are independent zero-mean Gaussian random variables with variances σ_ℓ given by (92). In order to find the integration limits we rewrite the matrix equation (99) as

$$q_\ell = t_{\ell+1} - \frac{1}{\ell}\sum_{i=1}^{\ell} t_i, \quad \ell = 1, 2, \ldots, m-1 \tag{100}$$

$$q_\ell = t_{\ell+2} - \frac{1}{\ell - m + 1}\sum_{i=m+1}^{\ell+1} t_i, \ \ell = m, \ldots, L-2 \tag{101}$$

$$q_{L-1} = bn\sum_{i=1}^{m} t_i + (\xi_{L-1} - \lambda_{\mathrm{A}1})\sum_{i=m+1}^{L} t_i \tag{102}$$

$$q_L = bn\sum_{i=1}^{m} t_i + (\xi_L - \lambda_{\mathrm{A}1})\sum_{i=m+1}^{L} t_i. \tag{103}$$

From the condition $\boldsymbol{t} \geq (\alpha \mathbf{1}_{(m)} \ \eta \mathbf{1}_{(n)})$ and (103) it follows directly that

$$q_L \geq bnm\alpha + n\eta(\xi_L - \lambda_{\mathrm{A}1}). \tag{104}$$

To estimate the limits for q_ℓ, $\ell = 1, 2, \ldots, m-1$, given by (100), we consider the linear combinations

$$q_L + \varphi q_\ell = \left(bn - \frac{\varphi}{\ell}\right)\sum_{i=1}^{\ell} t_i + (bn + \varphi)t_{\ell+1} + bn\sum_{i=\ell+2}^{m} t_i + (\xi_L - \lambda_{\mathrm{A}1})\sum_{i=m+1}^{L} t_i, \quad \ell = 1, 2, \ldots, m-1. \tag{105}$$

If $\varphi \in [-bn, bn\ell]$ then all the coefficients in (105) that multiply t_i, $i = 1, 2, \ldots, L$, are positive and their sum is equal to $bnm + n(\xi_L - \lambda_{\mathrm{A}1})$. Therefore we have

$$q_L + \varphi q_\ell \geq bnm\alpha + n\eta(\xi_L - \lambda_{\mathrm{A}1}), \quad \ell = 1, 2, \ldots, m-1.$$

The values $\varphi = -bn$ and $\varphi = bn\ell$ minimize the domain of integration. Thus we obtain that

$$-\frac{q_L - n(bm\alpha + \eta(\xi_L - \lambda_{\mathrm{A}1}))}{bn\ell} \leq q_\ell \leq \frac{q_L - n(bm\alpha + \eta(\xi_L - \lambda_{\mathrm{A}1}))}{bn}, \quad \ell = 1, \ldots, m-1. \tag{106}$$

Analogously, to estimate the limits for q_ℓ, $\ell = m, \ldots, L-2$, from (101), we consider linear combinations

$$q_L + \varphi q_\ell = bn\sum_{i=1}^{m} t_i + \left((\xi_L - \lambda_{\mathrm{A}1}) - \frac{\varphi}{\ell - m + 1}\right)\sum_{i=m+1}^{\ell+1} t_i + (\varphi + (\xi_L - \lambda_{\mathrm{A}1}))t_{\ell+2} + (\xi_L - \lambda_{\mathrm{A}1})\sum_{i=\ell+3}^{L} t_i$$

and conclude that for all values of φ in the interval $\varphi \in [-(\xi_L - \lambda_{\mathrm{A}1}),\ (\xi_L - \lambda_{\mathrm{A}1})(\ell - m - 1)]$, all the coefficients multiplying t_ℓ, $\ell = 1, 2, \ldots, L$ are positive and they sum up to $bnm + n(\xi_L - \lambda_{\mathrm{A}1})$. This implies that

$$q_L + \varphi q_\ell \geq bnm\alpha + n\eta(\xi_L - \lambda_{\mathrm{A}1}), \quad \ell = m, m+1, \ldots, L-2$$

from where it follows that for our set of φ, the shortest integration interval is

$$-\frac{q_L - n(bm\alpha + \eta(\xi_L - \lambda_{\mathrm{A}1})}{(\ell - m + 1)(\xi_L - \lambda_{\mathrm{A}1})} \leq q_\ell \leq \frac{q_L - n(bm\alpha + \eta(\xi_L - \lambda_{\mathrm{A}1}))}{(\xi_L - \lambda_{\mathrm{A}1})}, \quad \ell = m, \ldots, L-2. \tag{107}$$

Finally, we consider the linear combination of (102) and (103):

$$q_L + \varphi q_{L-1} = (1 + \varphi)bn\sum_{i=1}^{m} t_i + ((\xi_L - \lambda_{\mathrm{A}1}) + \varphi(\xi_{L-1} - \lambda_{\mathrm{A}1}))\sum_{i=m+1}^{L} t_i.$$

For $\varphi \in [-1,\ (\xi_L - \lambda_{\mathrm{A}1})/(\lambda_{\mathrm{A}1} - \xi_{L-1})]$ all coefficients on the right-hand side are positive and their sum is equal to $bnm(1+\varphi) + n((\xi_L - \lambda_{\mathrm{A}1}) + \varphi(\xi_{L-1} - \lambda_{\mathrm{A}1}))$. By taking $\varphi = -1$, from the condition that $\boldsymbol{t} \geq (\alpha \mathbf{1}_{(m)} \ \eta \mathbf{1}_{(n)})$, we obtain that

$$q_{L-1} \leq q_L - n\eta(\xi_L - \xi_{L-1}). \tag{108}$$

Similarly, for $\varphi = (\xi_L - \lambda_{\mathrm{A}1})/(\lambda_{\mathrm{A}1} - \xi_{L-1})$ we obtain

$$q_{L-1} \geq \frac{q_L(\xi_{L-1} - \lambda_{\mathrm{A}1}) + bmn\alpha(\xi_L - \xi_{L-1})}{\xi_L - \lambda_{\mathrm{A}1}}. \tag{109}$$

Inequalities (104), (106), (107), (108), and (109) define the integration limits over the probability density function $f(\boldsymbol{q})$ for obtaining an upper bound on the probability $\Pr(\boldsymbol{t} \geq \boldsymbol{\alpha})$. In order to have unit-variance variables with the $\mathcal{N}(0, 1)$ distribution, we introduce the following change of variables: $y_1 = q_L/\sqrt{\sigma_L}$, $y_2 = q_{L-1}/\sqrt{\sigma_{L-1}}$, and $x_\ell = q_\ell/\sqrt{\sigma_\ell}$, $\ell = 1, 2, \ldots, L-2$, which yields the expression (59) in Theorem 4. ■

REFERENCES

[1] P. Elias, "List decoding for noisy channels," in *Proc. IRE WESCON Conf. Record*, 1957, vol. 2, pp. 94–104.

[2] N. Seshadri and C.-E. W. Sundberg, "List Viterbi decoding algorithms with applications," *IEEE Trans. Comm.*, vol. 42, no. 2/3/4, pp. 313–323, 1994.

[3] F. K. Soong and E.-F. Huang, "A tree-trellis-based fast search for finding the N best sentence hypotheses in continuous speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc. (ICASSP)*, Toronto, Canada, Apr. 1991, vol. 1, pp. 705–708.

[4] M. Röder and R. Hamzaoui, "Fast tree-trellis list Viterbi decoding," *IEEE Trans. Comm.*, vol. 54, no. 3, pp. 453–461, Mar. 2006.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Intern. Conf. on Commun., ICC93*, Geneva, Switzerland, May 1993, vol. 2, pp. 1064 – 1070.

[6] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[7] P. A. Martin, D. P. Taylor, and M. C. Fossorier, "Soft-input soft-output list-based decoding algorithm," *IEEE Trans. Comm.*, vol. 52, no. 2, pp. 252–262, Feb. 2004.

[8] M. Lončar, R. Johannesson, I. E. Bocharova, and B. D. Kudryashov, "Iterative BEAST-decoding of product codes," in *Proc. 4th Intern. Symp. Turbo Codes & Related Topics*, Munich, Germany, 2006.

[9] D. J. Love, S. Hosur, A. Batra, and R. W. Heath Jr., "Space-time Chase decoding," *IEEE Trans. Wireless Comm.*, vol. 4, no. 5, pp. 2035–2039, Sept. 2005.

[10] J. Hagenauer and C. Kuhn, "Turbo equalization for channels with high memory using a list-sequential equalizer," in *Proc. Int. Symp. Turbo Codes*, ENST de Bretagne, France, Sept. 2003.

[11] V. V. Zyablov and M. S. Pinsker, "List cascade decoding," *Prob. Inform. Transm.*, vol. 17, no. 4, pp. 29–34, Oct.-Dec. 1982.

[12] V. Blinovsky, "Code bounds for finite-length-list decoding," *Problems of Inform. Transm.*, vol. 22, no. 1, pp. 7–14, 1986.

[13] P. Elias, "Error-correcting codes for list decoding," *IEEE Trans. Inform. Theory*, vol. 37, no. 1, pp. 5–12, Jan. 1991.

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

[15] V. Blinovsky, "Exponent of the probability of list-of-L decoding error at low rates," in *Proc. IEEE Int. Symp. Inform. Theory*, Washington, DC, USA, June 2001.

[16] V. Blinovsky, "Multiple packing of the Euclidean sphere," *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1334–1337, May 1999.

[17] E. R. Berlekamp, "The technology of error-correcting codes," *IEEE Trans. Inform. Theory*, vol. 68, no. 8, pp. 564–593, May 1980.

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

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

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

[21] R. M. Fano, *Transmission of Information*, M.I.T. Press, Cambridge, Massachusetts, 1961.

[22] R. G. Gallager, *Low-Density Parity-Check Codes*, Ph.D. thesis, MIT Press, Cambridge, MA, USA, 1963.

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

[24] S. Yousefi, "Gallager first bounding technique for the performance evaluation of maximum-likelihood decoded linear binary block codes," *IEE Proc. Commun.*, vol. 153, no. 3, pp. 317–332, June 2006.

[25] S. Yousefi and A. K. Khandani, "A new upper bound on the ML decoding error probability of linear binary block codes in AWGN interference," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3026–3036, Dec. 2004.

[26] A. Mehrabian and S. Yousefi, "A new upper bound on the ML decoding error probability of linear binary block codes in AWGN interference," in *Proc. Conf. on Inform. Sciences and Systems (CISS)*, Baltimore, USA, Mar. 2005.

[27] B. D. Kudryashov, "List decoding for the Gaussian noise channel," *Problems of Inform. Transm.*, vol. 27, no. 3, pp. 30–38, 1991.

[28] R. Herzog and C. Weiss, "Improved tight performance bounds on concatenated codes," in *Proc. Global Telecom. Conf. (GLOBECOM)*, Rio de Janeiro, Brazil, Dec. 1999.

[29] R. A. Horn and C. R. Johnson, *Matrix analysis*, University press, Cambridge, UK, 1985.

[30] V. Blinovsky, *Asymptotic Combinatorial Coding Theory*, Kluwer Academic Publishers, Norwell, MA, USA, 1987.

[31] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic press, UK, 1985.