Linear Programming-Based Optimization of the Distance Spectrum of Linear Block Codes

Gianluigi Ferrari, *Member, IEEE,* and Keith M. Chugg, *Member, IEEE*

***Abstract*—In this correspondence, we describe an approach for the identification of good distance spectra for *possibly existing* binary linear block codes based on linear programming and the MacWilliams–Delsarte identities. Specifically, the linear program is defined by an expression characterizing the performance of a potential code in terms of its distance spectrum and constraints imposed by the MacWilliams–Delsarte identities. Using the union bound to characterize performance, our results suggest that the best distance spectrum is not a function of signal-to-noise ratio (SNR) above the cutoff rate SNR and also suggest the existence of several unknown, good codes. Characterizing the performance using the maximum spectral error component of the union bound suggests spectral thinning with decreasing SNR.**

***Index Terms*—Distance spectrum, group codes, linear programming, optimization methods.**

I. Introduction

Assuming maximum-likelihood (ML) decoding, the *free distance* of a code (i.e., the minimum possible nonzero distance over all codewords) is a good indicator of performance at moderate to high signal-to-noise ratio (SNR). Maximizing the free distance was the traditional design goal [1] for codes until the emergence of *turbo-like* codes [2]–[5], which perform well at SNRs where the minimum distance is not the dominant characteristic of the code. Analysis of turbo-like codes has highlighted the importance of the entire distance spectrum rather than just the minimum distance [6]–[12].

With this appreciation, it is natural to inquire about the best distance spectrum of a code for a given rate, block length, and channel. While this is the goal of this correspondence, it is difficult to address this inquiry exactly because i) an exact expression of performance in terms of the distance spectrum is not available and ii) a necessary and sufficient condition for the existence of a code with a given distance spectrum is not known. Even if these two issues were surmounted, the resulting computational optimization problem would likely be intractable. Therefore, in this correspondence, we utilize one of two proxy expressions for performance and rely on the *MacWilliams–Delsarte identities* [1] to constrain the distance spectra. This allows the problem to be formulated as a linear programming problem that can be solved numerically. This does not result in an optimized code design, but rather in a candidate distance spectrum with best shape. As will be described, there may or may not exist a code with such spectrum.

Linear programming techniques and the MacWilliams–Delsarte identities have been applied in the past, but for different purposes than in this correspondence. In [13], a linear programming approach was considered, but it did not rely on the MacWilliams–Delsarte identities. These identities were instead utilized in [14], [15], where they were used to derive an upper bound on the number of codewords, in order to obtain bounds on the rate of a binary code. The MacWilliams–Delsarte identities were also used in [16], [17], where the proposed linear program aimed at proving the nonexistence of specific binary linear codes.

This correspondence is organized as follows. In Section II, we formalize the notion of distance spectrum and recall the MacWilliams–Delsarte identities. In Section III, we formulate a linear program with objective function given by the union bound on the codeword error rate (WER), and solve it for the cases of an additive white Gaussian noise (AWGN) channel and a perfectly interleaved Rayleigh flat-fading channel. In Section IV, we consider an alternative linear program, based on the dominant spectral error component of the union bound. Finally, Section V contains concluding remarks.

II. Distance Spectrum and MacWilliams–Delsarte Identities

Let $C = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ be a code comprising M codewords of length n (i.e., $\boldsymbol{x}_i = (x_{i,1}, \ldots, x_{i,n})$), where each symbol $x_{i,j}$ belongs to the binary alphabet and where $d(\boldsymbol{x}_\mu, \boldsymbol{x}_\nu) \geq d_f$ if $\mu \neq \nu$, where $d(\cdot, \cdot)$ denotes the Hamming distance. The quantity d_f is defined as the *minimum* or *free distance* of the code. For each $i \in \{0, 1, \ldots, n\}$, we define a_i to be the average number of codewords at distance i from a given codeword, i.e.,

$$a_i = \frac{1}{M} \cdot |\{(\mu, \nu): d(\boldsymbol{x}_\mu, \boldsymbol{x}_\nu) = i\}| \tag{1}$$

where $|\cdot|$ represents the cardinality of the corresponding set. We define $\boldsymbol{a} = (a_0, a_1, \ldots, a_n)$ as the *distance spectrum* of the code [15], with

$$a_0 + a_1 + \cdots + a_n = M. \tag{2}$$

If the code is linear, the distance spectrum $\boldsymbol{a}$ of the code reduces to its weight spectrum. The free distance of the code reduces in this case to the minimum possible nonzero weight of the codewords.

The MacWilliams–Delsarte identities [1], [18] are derived from the relations existing between a code and its dual code [19]. In particular, let x be an indeterminate and define b_j, $j \in \{0, 1, \ldots, n\}$, by the polynomial equality

$$\frac{1}{M} \sum_{i=0}^{n} a_i (1-x)^i (1+x)^{n-i} = \sum_{j=0}^{n} b_j x^j. \tag{3}$$

In general, b_j appears to have no natural combinatorial significance, but if $\{a_i\}$ is the weight spectrum of a binary linear block code C, then b_j equals the number of codewords of Hamming weight j in the dual code of C [20]. The equality (3) can be rewritten in terms of the weight enumerator $A(x) = \sum_{i=0}^{n} a_i x^i$ of the code and the weight enumerator $B(x) = \sum_{i=0}^{n} b_i x^i$ of its dual code as follows [1]:

$$B(x) = \frac{1}{M} (1+x)^n A\left[(1-x)/(1+x)\right]. \tag{4}$$

From (3), it is possible to derive [21]

$$b_j = \sum_{i=0}^{n} a_i K_j(i) \tag{5}$$

where $\{K_j(i)\}$ are the so-called Krawtchouk polynomials [15]. The coefficient of x^j in the polynomial $(1-x)^i(1+x)^{n-i}$ defines $K_j(i)$ as

$$K_j(i) = \underbrace{\sum_{k=0}^{i} \sum_{l=0}^{n-i}}_{k+l=j} \binom{i}{k} \binom{n-i}{l} (-1)^k \tag{6}$$

Manuscript received May 13, 2002; revised December 27, 2002. The material in this correspondence was presented in part at the 2nd Asian-European Workshop on Information Theory, Breisach, Germany, June 2002.

G. Ferrari is with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy (e-mail: gianluigi.ferrari@unipr.it).

K. M. Chugg is with the Communications Sciences Institute, Department of Electrical Engineering–Systems, University of Southern California, Los Angeles, CA 90089-2565 USA (chugg@usc.edu).

Communicated by C. Carlet, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2003.813483

where, for the sake of conciseness, the notation fails to recall the obvious fact that $K_j(i)$ also depends on n. Hence, the following general constraints hold for a generic binary linear block code:

$$\begin{cases} a_0 = 1 \\ a_1 = \cdots = a_{d_f-1} = 0 \\ a_i \geq 0,\ i \in \{d_f,\ d_f+1,\ \ldots,\ n\} \\ \sum_{i=0}^{n} a_i K_0(i) \geq 0 \\ \sum_{i=0}^{n} a_i K_1(i) \geq 0 \\ \cdots \\ \sum_{i=0}^{n} a_i K_n(i) \geq 0. \end{cases} \tag{7}$$

The linear constraints (7) are necessary to characterize a binary linear code, but may not be sufficient. It is an open question whether it is possible to find additional constraints valid for a generic linear code.

III. Union Bound-Based Linear Program

We consider ML decoding of linear block codes. In this case, a simple upper bound on the WER is given by a union bound over all possible codewords [22]. Assuming linearity, this upper bound, in terms of the distance spectrum, can be written as

$$P_W \leq \sum_{i=d_f}^{n} a_i P_2(i), \tag{8}$$

where $P_2(i)$ denotes the pairwise error probability (PEP) between the all-zero codeword and a codeword of weight i. This error probability depends on the considered transmission channel, the code rate R, the considered SNR, and the detection strategy. For the sake of notational simplicity, we do not explicitly indicate the dependence of the PEP on the code rate and the SNR. Using the upper bound in (8) as an *objective function* Γ and constraining the coefficients $\{a_i\}$ to satisfy the MacWilliams–Delsarte identities, we can formulate the following linear program:

$$\text{Minimize} \qquad \Gamma = \sum_{i=1}^{n} a_i P_2(i) \tag{9}$$

$$\text{subject to} \quad \begin{cases} \sum_{i=1}^{n} a_i K_0(i) \geq -K_0(0) \\ \sum_{i=1}^{n} a_i K_1(i) \geq -K_1(0) \\ \cdots \\ \sum_{i=1}^{n} a_i K_n(i) \geq -K_n(0) \\ \sum_{i=1}^{n} a_i = M - 1. \end{cases} \tag{10}$$

The description of the hypothetical weight spectrum $\boldsymbol{a}$ is completed[1] by stating that $a_0 = 1$. To completely specify the above linear program, the integer n and the expression of the PEP $P_2(i)$ have to be specified. Since the coefficients $P_2(i)$ in (9) are known, the objective function Γ is linear in the variables $\{a_i\}$.

[1]Additional constraints of the form $a_i = 0,\ i = 1,\ \ldots,\ \delta$ could be added for some integer δ. This would limit the search, embedded in the optimization procedure, among distance spectra with minimum distance δ. To avoid losing any generality, we consider $\delta = 1$.

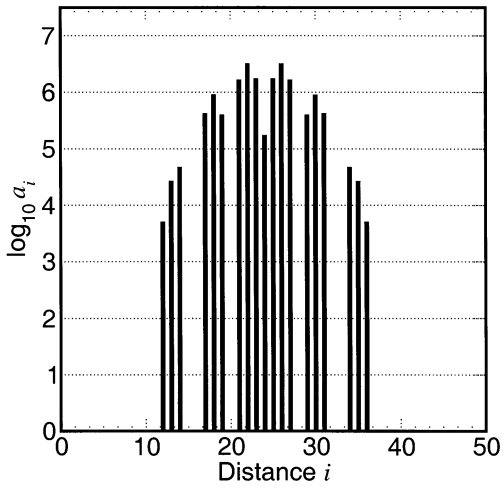

Fig. 1. Optimized distance spectrum, solution of a union bound-based linear program with symmetry condition, for $n = 48$ and $R = 1/2$ $(d_f = 12)$.

Solution of (9) and (10) is an integer programming problem and is, therefore, intractable for reasonably large-sized problems. If we relax the condition that the coefficients $\{a_i\}$ are integers and accept real values, we may use the *simplex method* [23] to solve (9) and (10) optimally within this context. Clearly, however, the resulting solution may not be the distance spectrum of a code because the best choice for $\{a_i\}$ may be real values. Even in the case that the optimizing values are integer, this only implies that the MacWilliams–Delsarte identities are satisfied by these integers, not that a corresponding code exists. Nonetheless, this approach should yield insight into the proper shape of the distance spectrum.

As mentioned earlier, the PEP depends on the particular transmission channel. In the following, we specialize and solve the linear program (9) and (10) in two cases.

A. AWGN Channel

Considering transmission over an AWGN channel with antipodal signaling, e.g., binary phase shift keying (BPSK), the PEP relative to soft-input decoding can be expressed as [22]

$$P_2(i) = Q\left(\sqrt{2\,\frac{E_S}{N_0}\,i}\right) = Q\left(\sqrt{2R\gamma_b i}\right) \tag{11}$$

where $Q(x)$ is the area under a mean zero, unit variance Gaussian density beyond x. Based on our experience, the linear program quickly becomes numerically intractable, for various reasons, for increasing codeword length n. In order to limit the number of independent variables $\{a_i\}$ we will consider relatively short block lengths and the following additional constrained *symmetry* of the distance spectrum:

$$\begin{cases} a_0 = a_n \\ a_1 = a_{n-1} \\ \cdots \\ a_{\lfloor \frac{n+1}{2} \rfloor - 1} = a_{\lceil \frac{n+1}{2} \rceil}. \end{cases} \tag{12}$$

The linear program in (9) and (10) can be straightforwardly modified to account for this additional constraint. A sufficient condition for the distance spectrum to be symmetric is that the all-1 codeword belongs to the codebook [21].

The solution of (9) and (10), imposing the symmetry condition and setting $n = 48$ and $R = 1/2$, is shown in Fig. 1. In this case, the minimum nonzero coefficient of the solution corresponds to $d_f = 12$. An interesting result is that the solution remains unchanged irrespective of

TABLE I
WEIGHT ENUMERATORS CORRESPONDING TO THE DISTANCE SPECTRA OBTAINED AS SOLUTIONS, IN THE INDICATED CASES, OF A UNION BOUND-BASED LINEAR PROGRAM WITH SYMMETRY CONDITION

Code	n	k	$A(x)$
Hamming	15	11	$1+x^{15}+35(x^3+x^{12})+105(x^4+x^{11})+168(x^5+x^{10})$ $+280(x^6+x^9)+435(x^7+x^8)$
Self-dual	16	8	$1+x^{16}+112(x^6+x^{10})+30x^8$
Golay	23	12	$1+x^{23}+253(x^7+x^{16})+506(x^8+x^{15})+1288(x^{11}+x^{12})$
Extended Golay	24	12	$1+x^{24}+759(x^8+x^{16})+2576x^{12}$
Hamming	31	26	$1+x^{31}+155(x^3+x^{28})+1085(x^4+x^{27})+5208(x^5+x^{26})22568(x^6+x^{25})$ $+82615(x^7+x^{24})+247845(x^8+x^{23})+628680(x^9+x^{22})+1383096(x^{10}+x^{21})$ $+2648919(x^{11}+x^{20})+4414865(x^{12}+x^{19})+6440560(x^{13}+x^{18})$ $+8280720(x^{14}+x^{17})+9398115(x^{15}+x^{16})$

the SNR considered. A general geometric justification of the independence of the obtained solution from the SNR comes from the following theorem. We assume that the scalar product of two vectors is denoted by the symbol $\cdot$, and that the all-1 vector is indicated with the symbol $\mathbf{1}=(1,\ldots,1)$.

Theorem 1: Let a linear program LP be intended to maximize[2] the objective function

$$\Gamma=\boldsymbol{a}\cdot\boldsymbol{f} \tag{13}$$

where $\boldsymbol{f}\in\mathbb{R}^n$, with respect to $\boldsymbol{a}\in\mathbb{R}^n$ subject to a number of constraints including

$$\boldsymbol{a}\in\mathbb{R}^+\times\mathbb{R}^+\times\cdots\times\mathbb{R}^+ \tag{14}$$

$$\boldsymbol{a}\cdot\mathbf{1}=a_1+a_2+\cdots+a_n=M \tag{15}$$

where $M\in\mathbb{R}$. The other constraints on $\boldsymbol{a}$ need not be specified. Let us assume that LP admits the solution $\boldsymbol{a}^{\max}$. Let $\mathcal{P}=\{\boldsymbol{x}\in\mathbb{R}^n\colon \boldsymbol{x}\cdot\mathbf{1}=0\}$ be the hyperplane passing through the origin and parallel to $\mathcal{P}'=\{\boldsymbol{x}\in\mathbb{R}^n\colon \boldsymbol{x}\cdot\mathbf{1}=M\}$.

Let LP1 be another linear program, the objective function of which is

$$\Gamma_1=\boldsymbol{a}\cdot\boldsymbol{f}_1 \tag{16}$$

with $\boldsymbol{f}_1\neq\boldsymbol{f}$ and where the set of constraints is exactly the same as in LP. Let $\boldsymbol{f}^{\mathcal{P}}(\boldsymbol{f}_1^{\mathcal{P}})$ be the projection of $\boldsymbol{f}(\boldsymbol{f}_1)$ onto $\mathcal{P}$. If $\boldsymbol{f}_1^{\mathcal{P}}=c\boldsymbol{f}^{\mathcal{P}}$, $c\in\mathbb{R}^+$, then LP1 admits $\boldsymbol{a}^{\max}$ as its solution.

Proof: See the Appendix. □

Considering the linear program (9) and (10) for a specific SNR, we can interpret it as a particular case of the linear program LP of Theorem 1, with f_i given by $P_2(i)$. Note that $P_2(1)>P_2(2)>\cdots>P_2(n)>0$, i.e., $P_2(i)$ is monotonically decreasing on i. Changing the SNR simply changes the rate of decrease, with respect to i, of $P_2(i)$. Above the SNR corresponding to the cutoff rate, the sum of these terms converges so that the rate of decrease with respect to i is very large (i.e., exponential). Hence, the family of vectors $\{(P_2(1),\ldots,P_2(n))\}$ corresponding to different SNRs are contained in a small cone in $\mathbb{R}^n$ that does not contain the vector $\mathbf{1}$ orthogonal to $\mathcal{P}$. In fact, this implies that the corresponding vectors obtained by projection onto the hyperplane $\mathcal{P}$—as explained in Theorem 1—lie near each other and have the same orientation (in the sense that the scalar product between any two of them is positive). It is then arguable that the solution of the linear program should not change significantly when changing the SNR. Numerically, we find that the solution does not depend on the SNR at all.

The solution of the linear program shown in Fig. 1 is not a feasible weight spectrum for a binary linear block code, in the sense that the coefficients $\{a_i\}$ are not integers. However, in several cases, the linear program has a solution where the coefficients $\{a_i\}$ are, within tight numerical precision limits, nonnegative integers. Obviously, when the code rate R approaches 1, the solution approaches a binomial distribution. In the limit $(R=1)$, it has to be binomial, since all n-tuples are possible codewords. Moreover, in several special cases the solutions are the distance spectra of well-known group codes. In Table I, we report some of the obtained distance spectra which exactly match with those of known group codes.[3] In these cases, we can conclude that these codes are "optimal" in the sense of minimizing the union bound (9) on the WER. In other cases, the solution of the linear program is a vector with nonnegative integers which does not correspond, to the best of our knowledge, to the distance spectrum of any known block code. This could predict the existence of still unknown codes. For example, solving the considered linear program for $n=34$ and $k=15$, the following weight enumerator is obtained:

$$\begin{aligned}A(x)=1&+x^{34}+374(x^{10}+x^{24})+1152(x^{11}+x^{23})\\&+1184(x^{12}+x^{22})+1848(x^{14}+x^{20})\\&+7040(x^{15}+x^{19})+4785(x^{16}+x^{18}).\end{aligned} \tag{17}$$

As one can see, the obtained distance spectrum would correspond to a $(34,15)$ binary linear code with $d_f=10$—provided that a search procedure found a code with such a weight spectrum. To the best of our knowledge, the $(34,15)$ group codes known up to now have at most $d_f=9$ [25]. Another example is the following distance spectrum,

[2]The theorem holds in the same way in the case of minimization. The proof can be straightforwardly modified.

[3]Incidentally, we observe that these codes are perfect or close to perfect.

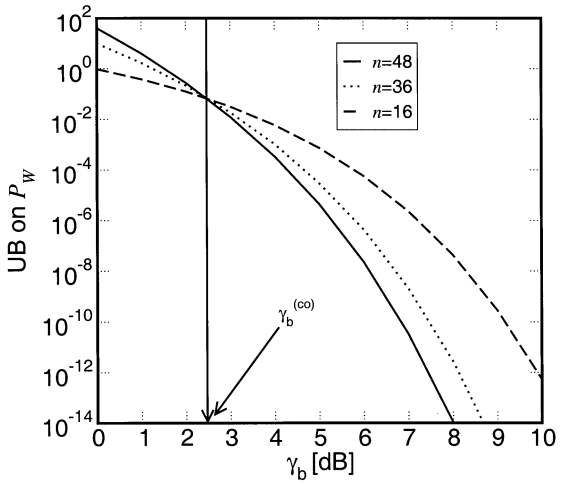

Fig. 2. Upper bounds on the WER over an AWGN channel relative to optimized distance spectra, solutions of a union bound-based linear program with symmetry condition, for $R = 1/2$ and increasing codeword length n. The curves cross at the SNR $\gamma_b^{(co)} = 2.46$ dB corresponding to a cutoff rate $R_0 = R = 1/2$ over an AWGN channel with BPSK.

obtained by solving the linear program with $n = 37$ and $k = 9$. In this case, the weight enumerator corresponding to the solution is

$$A(x) = 1 + x^{37} + 72(x^{15} + x^{22}) + 183(x^{16} + x^{21}). \tag{18}$$

The hypothetical code corresponding to (18) has $d_f = 15$. This achieves the highest possible allowed minimum distance [16]. In [26], a $(37, 9)$ code with maximum allowed $d_f = 15$ is proposed, and its distance spectrum is

$$A(x) = 1 + 91x^{15} + 125x^{16} + 107x^{19} + 96x^{20} + 57x^{23} + 34x^{34} + x^{37}. \tag{19}$$

We finally evaluated the upper bounds on the WER (9) corresponding to the obtained optimized distance spectra. In Fig. 2, the upper bounds relative to the case $R = 1/2$ are shown for increasing codeword length n. They cross at a specific SNR $\gamma_b^{(co)}$, such that the *cutoff rate* R_0 coincides with the code rate R. In the case of BPSK transmission over an AWGN channel, this SNR can be written [22] as

$$\gamma_b^{(co)} = \frac{1}{R} \ln\left[\frac{1}{2^{1-R} - 1}\right]. \tag{20}$$

Substituting for $R = 1/2$ in (20), $\gamma_b^{(co)} = 2.46$ dB is obtained, which is the SNR corresponding to the intersection of the various bounds in Fig. 2. This means that the proposed linear program leads to distance spectra which asymptotically (for increasing n) show decreasing error probability for any SNR such that $R < R_0$. This is reasonable, since the objective function is a union bound, i.e., a sort of averaged performance measure converging above the channel cutoff rate SNR [27]. This result does suggest that, above the cutoff rate SNR, the best distance spectrum is not dependent on the SNR.

B. Perfectly Interleaved Rayleigh Flat-Fading Channel

In this case, the general setting of the linear program does not change with respect to the AWGN channel case. The only difference lies on the expression of the PEP. For BPSK signaling over a perfectly interleaved Rayleigh flat-fading channel and in the case of coherent hard detection followed by decoding, the PEP can be written [22] as

$$P_2(i) = p^i \sum_{k=0}^{i-1} \binom{i-1+k}{k} (1-p)^k \tag{21}$$

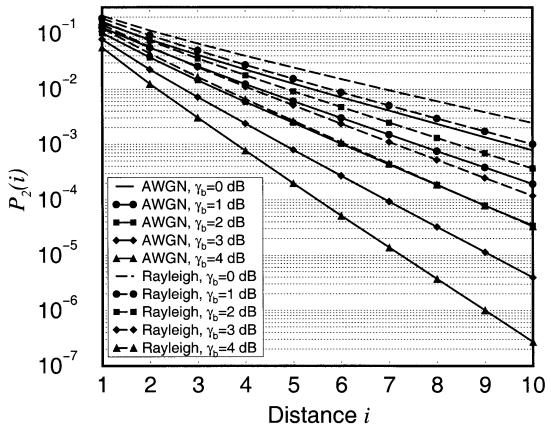

Fig. 3. Comparison between the PEP over an AWGN channel and the PEP over a Rayleigh flat-fading channel, as a function of the distance i, for $R = 1/2$ and increasing SNR γ_b.

where

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{R\gamma_b}{1 + R\gamma_b}}\right). \tag{22}$$

In order to justify the results obtained in the case of a Rayleigh flat-fading channel, we make a preliminary comparison between the PEP in the AWGN channel case and the PEP in the Rayleigh flat-fading channel case. In Fig. 3, the PEPs in the two cases, as functions of the distance, are compared for different SNRs. As one can see, in logarithmic scale the PEPs in the two cases are approximately linear, and the approximation is the better, the larger the distance or the higher the SNR. According to the given interpretation of Theorem 1, the distance spectrum solution of the linear program is expected to remain almost unchanged at any SNR. Moreover, Theorem 1 predicts that the distance spectrum obtained in the current case should be very similar to that obtained in the AWGN channel case. In fact, the PEP (21) in the Rayleigh flat-fading channel case is rapidly decreasing as in the AWGN channel case. Numerical results show that the solution of the linear program in the two channel cases—when setting in the same way all other parameters—does not change. The solution is exactly the same whether an AWGN channel or a perfectly interleaved Rayleigh flat-fading channel is considered.

In this case as well, we evaluated the upper bounds on the WER obtained with the optimized distance spectra. In Fig. 4, the upper bounds are shown in the case of $R = 1/2$ for increasing codeword length n. The same behavior as in the AWGN channel case is observed in the current case: the curves corresponding to different codeword lengths cross at a precise point, relative to the channel cutoff rate. In fact, in the case of BPSK $(q = 2)$ transmission over a Rayleigh flat-fading channel, the SNR corresponding to the channel cutoff rate can be expressed [22] as

$$\gamma_b^{(co)} = \frac{q(2^R - 1)}{(q - 2^R)R} = \frac{(2^R - 1)}{(1 - 2^{R-1})R}. \tag{23}$$

Substituting for $R = 1/2$, we find $\gamma_b^{(co)} = 4.51$ dB, which is exactly the SNR where the curves cross.

IV. Maximum Spectral Error Component-Based Linear Program

Based on the evidence available from existing turbo-like code designs, the insensitivity of the best distance spectrum to SNR is most likely due to the use of the union bound. To obtain a useful result for SNR below the cutoff rate SNR, another measure of performance is

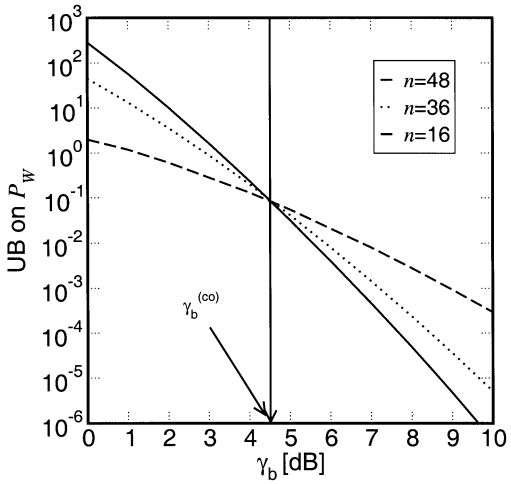

Fig. 4. Upper bounds on the WER over a Rayleigh flat-fading channel with optimized distance spectra, solutions of a union bound-based linear program with symmetry condition, for $R = 1/2$, and increasing codeword length n. The curves cross at the SNR $\gamma_b^{(co)} = 4.51$ dB corresponding to a cutoff rate $R_0 = R = 1/2$ over a Rayleigh flat-fading channel with BPSC and coherent detection.

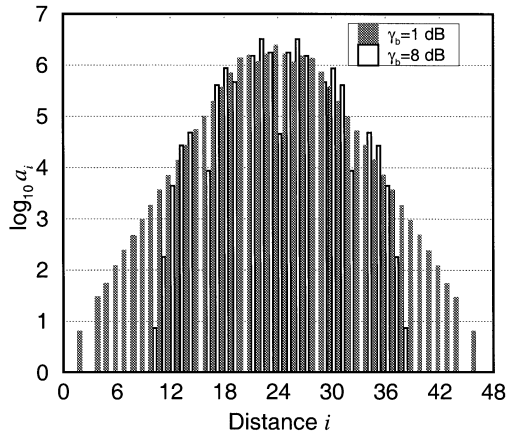

Fig. 5. Distance spectra, solutions of the maximum spectral error component-based linear program with symmetry condition, in the case with $n = 48$ and $R = 1/2$, for $\gamma_b = 1$ dB and $\gamma_b = 8$ dB.

required. However, adopting a complex bound will substantially complicate the optimization problem (see Section V).

In this section, we consider characterizing performance by the maximum term in the union bound. This is not a strict bound in any sense, but it is reasonable to expect that at lower SNRs this would more accurately track the exact WER. The resulting nonlinear optimization problem is

$$\text{Minimize } \max_i \{a_i P_2(i)\} \tag{24}$$

where the minimization is carried out over all possible $\boldsymbol{a}$, subject to the constraints (10). We refer to a term $a_i P_2(i)$ as a *spectral error component*. It is possible to transform this nonlinear program into an equivalent linear program by defining G as

$$G = \max_i \{a_i P_2(i)\} \tag{25}$$

and noting that

$$G \geq a_i P_2(i), \qquad i \in \{1, \ldots, n\}. \tag{26}$$

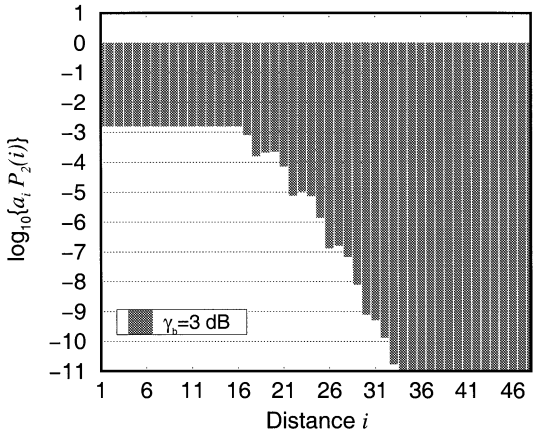

Fig. 6. Distribution of the spectral error components in the AWGN case, relative to the distance spectrum solution of the maximum spectral error component-based linear program with symmetry condition, in the case with $n = 48$ and $R = 1/2$, for $\gamma_b = 3$ dB.

Considering the set of constraints given by (26) and combining them with the constraints (10), the following linear program can be formulated:

$$\text{Minimize} \qquad G = \sum_{i=1}^{n} 0 \cdot a_i + 1 \cdot G \tag{27}$$

$$\text{subject to} \quad \begin{cases} G \geq 0 \\ G \geq a_1 P_2(1) \\ \cdots \\ G \geq a_n P_2(n) \\ \sum_{i=1}^{n} a_i K_0(i) \geq -K_0(0) \\ \cdots \\ \sum_{i=1}^{n} a_i K_n(i) \geq -K_n(0) \\ \sum_{i=1}^{n} a_i = M - 1. \end{cases} \tag{28}$$

There is a fundamental difference with respect to the linear program in (9) and (10). In (27) and (28), the PEP appears in the constraints (28), but does not in the objective function (27). This means that, in the linear program (27) and (28), changing the SNR leaves unmodified the objective function, but modifies the domain where $\boldsymbol{a}$ can vary. In this case, we expect that the optimized distance spectrum will depend on the particular SNR. Moreover, since for large SNR it holds that

$$\sum_{i=1}^{n} a_i P_2(i) \approx \max_{i \in \{1,\ldots,n\}} \{a_i P_2(i)\} \tag{29}$$

we expect that solving (27) and (28) and (9) and (10) will yield similar results at moderate to high SNR.

In order to make a comparison with the results obtained in Section III, we apply the symmetry condition to the linear program (27) and (28). For $n = 48$ and $R = 1/2$, the optimized distance spectra, for $\gamma_b = 1$ dB and $\gamma_b = 8$ dB, are shown in Fig. 5. As one can see, for $\gamma_b = 8$ dB, the solution is very similar to the distribution in Fig. 1. However, there is a significant difference at low SNR, where the tails of the optimized distribution in the current case become significant. In Fig. 6, the spectral error components corresponding to the distance spectrum obtained for $\gamma_b = 3$ dB (around the cutoff rate) are shown.

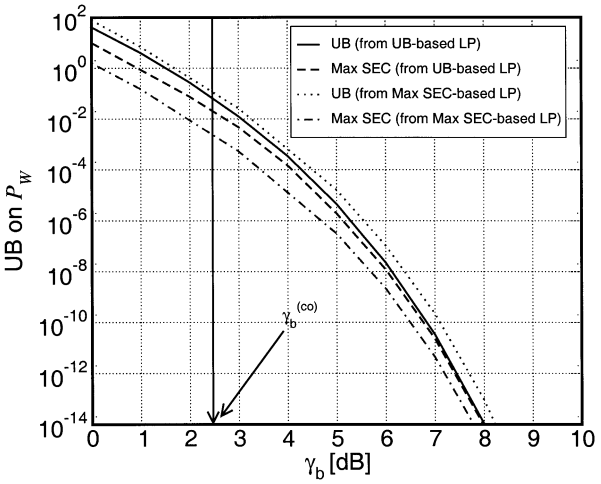

Fig. 7. Comparison, for $n = 48$ and $R = 1/2$, between performance curves relative to the union bound-based linear program (UB) and curves relative to the maximum spectral error component-based linear program (max SEC). In particular, in each case two curves are shown, corresponding to the union bound and the maximum spectral error component, respectively.

The spectral error components relative to the indexes between 1 and 16 have almost the same (maximum) value.

In order to further compare the results obtained with the two proposed linear programs, the maximum spectral error component, corresponding to the solution of the union bound-based linear program, and the union bound, corresponding to the solution of the maximum spectral error component-based linear program are also evaluated. The results are shown in Fig. 7. Even if the solution of the union bound-based linear program does not depend on the SNR, the index $i_{\max}$ of the maximum spectral error component does. While $i_{\max} = 18$ for $\gamma_b = 0$ dB, the index $i_{\max}$ reduces to $d_f = 12$ for $\gamma_b \geq 6$ dB. In Fig. 7, the curves obtained with the maximum spectral error component-based linear program are such that for any SNR, the corresponding solution of the linear program is used. In particular, for $\gamma_b = 1$ dB the spectral error components have almost the same maximum value for all indexes between 2 and 20 $(a_1 = 0)$, while for $\gamma_b = 7$ dB, the maximum value approximately corresponds to the indexes between 1 and 12. This justifies the fact that the union bound derived from the maximum spectral error component-based linear program converges to the union bound derived from the union bound-based linear program only for very large SNR.

V. Concluding Remarks

In this correspondence, we considered a linear programming approach to the optimization of the weight spectrum of a hypothetical binary linear block code transmitted over an AWGN channel or a perfectly interleaved Rayleigh flat-fading channel. Linear optimization was carried out, where the objective function was a union bound on the WER, or the maximum spectral error component of this bound, based on a hypothetical weight spectrum. The constraints were derived from the MacWilliams–Delsarte identities. With the union bound-based approach, in several cases the obtained distance spectra corresponded to those of known group codes, and in other cases, the possible existence of new group codes was suggested, but not proven. The union bound-based approach led to the identification of distance spectra which show asymptotically (for increasing codeword length) decreasing error probability for an SNR above that corresponding to the channel cutoff rate. Ironically, although motivated by the belief that the entire distance spectrum should eventually determine the performance of a code, our results suggest that above the cutoff rate SNR the *minimum distance* is the primary factor. This is most likely due to the use of the union bound and, in fact, the maximum spectral error component-based linear program suggests that the conclusions might be different in the SNR region below the channel cutoff rate SNR. Specifically, the results using this performance metric suggest that the best design should sacrifice a smaller free distance for a thinner spectrum.

Extension of the proposed approach to the low-SNR region in a meaningful manner is desirable. Substituting for a tighter upper bound, as for example the tangential sphere bound [28], valid for SNRs below that corresponding to the channel cutoff rate, could lead to the identification of distance spectra which guarantee asymptotically low error probability near the channel capacity [29]. However, our previous attempts to incorporate the tangential sphere bound have not been fruitful. While this is a conceptually simple modification of the union bound-based approach, in practice, we have found that the optimization problem becomes extremely difficult. Furthermore, since the union bound results in a linear program, we can ensure that the stated problem is solved optimally [24], while a tangential sphere bound will result in a nonlinear program where such claims cannot be made.

An open problem, from a numerical point of view, is the extension of the proposed method to larger codeword lengths and to nonsymmetric weight spectra. This could be accomplished by considering more sophisticated tools of linear optimization. Moreover, in the cases where the solution of the linear programs (9) and (10) or (27) and (28) is constituted by noninteger numbers, it would be interesting to look at the nearest (according to a suitable measure) integer distribution, i.e., a possible weight spectrum.

Appendix
Proof of Theorem 1

Before proving Theorem 1, we present two lemmas which are useful. We recall the notation used in the previous sections. In particular, we denote by $\boldsymbol{a} \in \mathbb{R}^n$ an n-dimensional vector $(a_1, a_2, \ldots, a_n)$. Given two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we define their scalar product as

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + \cdots + a_n b_n \in \mathbb{R}.$$

Lemma 1: Let us consider two vectors $\boldsymbol{a}, \boldsymbol{f} \in \mathbb{R}^n$ and the hyperplane $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \cdot \boldsymbol{1} = 0\}$ ($\mathcal{P}$ passes through the origin). Let us denote by $\boldsymbol{f}^{\mathcal{P}}$ the projection of $\boldsymbol{f}$ onto the hyperplane $\mathcal{P}$. If $\boldsymbol{a}' = \boldsymbol{a} + c\boldsymbol{f}^{\mathcal{P}}$, $c \in \mathbb{R}^+$, then

$$\boldsymbol{a}' \cdot \boldsymbol{f} \geq \boldsymbol{a} \cdot \boldsymbol{f}. \tag{30}$$

Proof: By linearity of the scalar product, we can write

$$\boldsymbol{a}' \cdot \boldsymbol{f} = (\boldsymbol{a} + c\boldsymbol{f}^{\mathcal{P}}) \cdot \boldsymbol{f} = \boldsymbol{a} \cdot \boldsymbol{f} + c\boldsymbol{f}^{\mathcal{P}} \cdot \boldsymbol{f}. \tag{31}$$

From the definition of projection, it follows $\boldsymbol{f}^{\mathcal{P}} \cdot \boldsymbol{f} \geq 0$. This proves (30). □

We interpret Lemma 1 as follows. Let us define the hyperplane

$$\mathcal{P}' = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \boldsymbol{y} + \boldsymbol{a}, \forall \boldsymbol{y} \in \mathcal{P}\}$$

i.e., the hyperplane parallel to $\mathcal{P}$ which contains the vector $\boldsymbol{a}$. For the sake of notational simplicity, we will use the notation $\mathcal{P}' = \mathcal{P} + \boldsymbol{a}$. Considering the scalar product $\boldsymbol{a} \cdot \boldsymbol{f}$ (a hypothetical objective function), if $\boldsymbol{a}$ "moves" inside $\mathcal{P}'$ in the direction of $\boldsymbol{f}^{\mathcal{P}}$, then the hypothetical objective function increases. Lemma 1 can be generalized as follows.

Lemma 2: Let us consider $\boldsymbol{a}, \boldsymbol{f}, \boldsymbol{f}^{\mathcal{P}} \in \mathbb{R}^n$ and the hyperplane $\mathcal{P}$ as defined in Lemma 1. Let us consider a vector $\boldsymbol{g} \in \mathcal{P}$ such that $\boldsymbol{g} \cdot \boldsymbol{f}^{\mathcal{P}} \geq 0$. If $\boldsymbol{a}'' = \boldsymbol{a} + \boldsymbol{g}$, then

$$\boldsymbol{a}'' \cdot \boldsymbol{f} \geq \boldsymbol{a} \cdot \boldsymbol{f}. \tag{32}$$

Proof: Owing to the linearity of the scalar product, in order to prove (32) it is sufficient to prove that $\boldsymbol{g} \cdot \boldsymbol{f} \geq 0$. Two cases can be distinguished.

1) $\boldsymbol{f}$ is orthogonal to $\mathcal{P}$. This case is trivial. In fact, any vector $\boldsymbol{g} \in \mathcal{P}$ is orthogonal to $\boldsymbol{f}$, i.e., $\boldsymbol{g} \cdot \boldsymbol{f} = 0$.
2) $\boldsymbol{f}$ is not orthogonal to $\mathcal{P}$. In this case, $\boldsymbol{f} \cdot \boldsymbol{f}^{\mathcal{P}} > 0$. The hyperplane

$$\mathcal{F} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = c_1\ \boldsymbol{f} + c_2 \boldsymbol{f}^{\mathcal{P}},\ \forall\, c_1,\ c_2 \in \mathbb{R}\}$$

is orthogonal to $\mathcal{P}$. Moreover,

$$\mathcal{F} \cap \mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = r\boldsymbol{f}^{\mathcal{P}},\ \forall\, r \in \mathbb{R}\} = l$$

is a hyperline inside $\mathcal{P}$. It follows that if $\boldsymbol{g} \cdot \boldsymbol{f}^{\mathcal{P}} = 0$, then $\boldsymbol{g}$ is orthogonal to $\mathcal{F}$ so $\boldsymbol{g} \cdot \boldsymbol{f} = 0$. If $\boldsymbol{g} \cdot \boldsymbol{f}^{\mathcal{P}} > 0$, then $\boldsymbol{g}$ is not orthogonal to $\mathcal{F}$. Then $\boldsymbol{f} \cdot \boldsymbol{f}^{\mathcal{P}} > 0$ results in $\boldsymbol{g} \cdot \boldsymbol{f} > 0$. □

Lemma 2 can be interpreted as follows. Given the hyperplane $\mathcal{P}$ and the vector $\boldsymbol{f}^{\mathcal{P}}$ as defined in Lemma 1, one can consider the hyperline $l \subset \mathcal{P}$ orthogonal to $\boldsymbol{f}^{\mathcal{P}}$. The hyperplane $\mathcal{P}$ is divided into two regions by l. We define $\mathcal{R}_1$ as the region (including the hyperline l) which contains $\boldsymbol{f}^{\mathcal{P}}$, and $\mathcal{R}_2 = \mathcal{P} \backslash \mathcal{R}_1$ the complementary region. Considering the hyperplane $\mathcal{P}' = \mathcal{P} + \boldsymbol{a}$, it is possible to identify two regions, $\mathcal{R}_1' \subset \mathcal{P}'$ and $\mathcal{R}_2' \subset \mathcal{P}'$, such that $\mathcal{R}_1' = \mathcal{R}_1 + \boldsymbol{a}$ and $\mathcal{R}_2' = \mathcal{R}_2 + \boldsymbol{a}$. A generic vector $\boldsymbol{a}''$ as defined in Lemma 2 is such that $\boldsymbol{a}'' \in \mathcal{R}_1'$.

We now propose a proof of Theorem 1.

Proof [Theorem 1]: The solution $\boldsymbol{a}^{\max}$ of the LP must be such that

$$\boldsymbol{a}^{\max} \in \mathcal{P}' = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \cdot \mathbf{1} = M,\ M \in \mathbb{R}\}.$$

Let us consider the hyperline $l \subset \mathcal{P}$ and the regions $\mathcal{R}_1 \subset \mathcal{P}$ and $\mathcal{R}_2 \subset \mathcal{P}$ as defined above. We define by $l' = l + \boldsymbol{a}^{\max} \subset \mathcal{P}'$ the hyperline parallel to l such that $\boldsymbol{a}^{\max} \in l'$. Similarly, we consider $\mathcal{R}_1' = \mathcal{R}_1 + \boldsymbol{a}^{\max}$ and $\mathcal{R}_2' = \mathcal{R}_2 + \boldsymbol{a}^{\max}$. Obviously, $l' \subset \mathcal{R}_1'$. Since we assumed that LP admits a solution, it follows that $\mathcal{R}_1' \cap \mathcal{D} = \boldsymbol{a}^{\max}$, where $\mathcal{D}$ is the domain determined by all the constraints of LP (including (14) and (15)). In fact

a) if $(l' \cap \mathcal{D}) \backslash \{\boldsymbol{a}^{\max}\} \neq \{\emptyset\}$, then, considering a vector $\boldsymbol{a}'' \in l' \cap \mathcal{D}$, it follows that $\boldsymbol{a}'' \cdot \boldsymbol{f} = \boldsymbol{a}^{\max} \cdot \boldsymbol{f}$. According to the *fundamental theorem of optimization* [24], the solution of a linear program, when it exists, is unique. Since we assumed that LP admits a solution, it must be $l' \cap \mathcal{D} = \boldsymbol{a}^{\max}$;

b) if $(\mathcal{R}_1' \backslash l') \cap \mathcal{D} \neq \{\emptyset\}$, then according to Lemma 2 it is possible to find a vector $\boldsymbol{a}'' \in \mathcal{R}_1' \backslash l'$, $\boldsymbol{a}'' \neq \boldsymbol{a}^{\max}$, such that $\boldsymbol{a}'' \cdot \boldsymbol{f} > \boldsymbol{a}^{\max} \cdot \boldsymbol{f}$. This contradicts the assumption that $\boldsymbol{a}^{\max}$ is the solution of LP, i.e., that it maximizes the scalar product (13). It follows that $(\mathcal{R}_1' \backslash l') \cap \mathcal{D} = \{\emptyset\}$.

From a) and b) it follows that $\mathcal{R}_1' \cap \mathcal{D} = \boldsymbol{a}^{\max}$. Let us now consider the linear program LP1. It is possible to prove that

$$\boldsymbol{a}^{\max} \cdot \boldsymbol{f}_1 = \max_{\boldsymbol{a}} \in \mathcal{D}\{\boldsymbol{a} \cdot \boldsymbol{f}_1\}.$$

In fact, according to Lemma 2, $\boldsymbol{a}'' \cdot \boldsymbol{f}_1 \geq \boldsymbol{a}^{\max} \cdot \boldsymbol{f}_1$ only if $\boldsymbol{a}'' \in \mathcal{R}_1' \backslash \{\boldsymbol{a}^{\max}\}$. However, we have shown above that $(\mathcal{R}_1' \backslash \{\boldsymbol{a}^{\max}\}) \cap \mathcal{D} = \{\emptyset\}$. It thus follows that the solution of LP1 must be $\boldsymbol{a}^{\max}$. □

ACKNOWLEDGMENT

Prof. L. R. Welch and Prof. P. V. Kumar, from the Communication Sciences Institute, University of Southern California, are kindly acknowledged for important and very helpful discussions. The authors would also like to thank the Editor and the Reviewers, in particular Reviewer 1, for their encouragement and very detailed, precise comments.

REFERENCES

[1] G. C. Clark and J. B. Cain, *Error Correction Coding for Digital Communications*. New York: Plenum, 1988.
[2] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
[3] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.
[4] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 909–926, May 1998.
[5] L. Ping, X. Huang, and N. Phamdo, "Zigzag codes and concatenated zigzag codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 800–807, Feb. 2001.
[6] R. Podemski, W. Holubowicz, C. Berrou, and A. Glavieux, "Distance spectrum of turbo codes," in *Proc. IEEE Symp. Information Theory*, June 1995, p. 34.
[7] Y. V. Svirid, "Weight distribution of turbo-codes," in *Proc. IEEE Symp. Information Theory*, June 1995, p. 38.
[8] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," Jet Propulsion Labs., Pasadena, CA, Tech. Rep. TDA Progress Report 42-122, Aug. 1995.
[9] O. Y. Takeshita, M. P. C. Fossorier, and D. J. Costello, "A new technique for computing the weight spectrum of turbo-codes," *IEEE Commun. Lett.*, vol. 3, pp. 251–253, Aug. 1999.
[10] G. Battail, "A conceptual framework for understanding turbo codes," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 245–254, Feb. 1998.
[11] ——, "Construction explicite de bons codes longs," *Ann. Telecommun.*, vol. 44, pp. 392–404, July/Aug. 1989.
[12] L. Perez, J. Seghers, and D. J. Costello, "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1698–1709, Nov. 1996.
[13] E. J. McCluskey , Jr., "Error-correcting codes—A linear programming approach," *Bell Syst. Tech. J.*, vol. 38, pp. 1485–1512, Nov. 1959.
[14] L. R. Welch, R. J. McEliece, and H. Rumsey, "A low-rate improvement on the Elias bound," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 676–678, Sept. 1974.
[15] R. J. McEliece, E. R. Rodemich, H. Rumsey, and L. R. Welch, "New upper bounds on the rate of a code via Delsarte–MacWilliams inequalities," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 157–166, Mar. 1977.
[16] R. Hill and K. L. Traynor, "The nonexistence of certain binary linear codes," *IEEE Trans. Inform. Theory*, vol. 36, pp. 917–922, July 1990.
[17] A. E. Brouwer, "The linear programming bound for binary linear codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 677–680, Mar. 1993.
[18] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.
[19] P. Delsarte and I. Levenshtein, "Association schemes and coding theory," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2477–2504, Oct. 1998.
[20] P. Delsarte, "Bounds for unrestricted codes, by linear programming," Philips Res. Rep., Eindhoven, The Netherlands, Tech. Rep., 1972.
[21] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. New York: North-Holland, 1977.
[22] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[23] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.
[24] D. G. Luenberger, *Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley, 1984.
[25] P. Farkaš and S. Herrera-Garcia, "Three new optimal [34,15,9] codes," *Electronics Letters.com (web journal)*, Nov. 2001.
[26] P. Farkaš and K. Brühl, "Three best binary linear block codes of minimum distance fifteen," *IEEE Trans. Inform. Theory*, vol. 40, pp. 949–951, May 1994.
[27] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[28] G. Poltyrev, "Bounds on the decoding of error probability of binary linear codes via their spectra," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1284–1292, July 1994.
[29] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, pp. 379–423, July 1948.