Multidimensional Coded Modulation in Block-Fading Channels

Albert Guillén i Fàbregas, *Member, IEEE*, and Giuseppe Caire, *Fellow, IEEE*

***Abstract*—We study coded modulation over multidimensional signal sets in Nakagami-m block-fading channels. We consider the optimal diversity reliability exponent of the error probability when the multidimensional constellation is obtained as the rotation of complex-plane signal constellations. We show that multidimensional rotations of full dimension achieve the optimal diversity reliability exponent, also achieved by Gaussian constellations. Rotations of full dimension induce a large decoding complexity, and in some cases it might be beneficial to use multiple rotations of smaller dimension. We also study the diversity reliability exponent in this case, which yields the optimal rate–diversity–complexity tradeoff in block-fading channels with discrete inputs.**

***Index Terms*—Block-fading channels, diversity, linear rotations, maximum distance-separable (MDS) codes, outage probability.**

I. Introduction

Rotated multidimensional constellations were proposed as a way of achieving high reliability with *uncoded* modulation in fading channels. Rotated constellations have been extensively studied, and have been shown to be an effective technique to achieve full-rate and full-diversity transmission in scalar and multiple-input multiple-output (MIMO) fading channels (see [1] and references therein).

In this work, we study the problem of constructing coded modulation schemes over multidimensional signal sets, obtained by rotating complex-plane signal constellations, for block-fading channels with B fading blocks per codeword [2]. The block-fading channel is a useful model for transmission over slowly varying fading channels, such as orthogonal frequency division multiplexing (OFDM) or slow time-frequency-hopped systems such as Global System for Mobile Communication (GSM) or Enhanced Data GSM Environment (EDGE).

Full-diversity rotations of dimension B induce large decoding complexity since the set of candidate points for detection at a given time instant is exponential with B. Uncoded rotations are typically decoded with the sphere decoder [3] to avoid exhaustive search. When coded modulation is used, the code itself can help to achieve full diversity. This means that sometimes rotations of smaller dimension $N < B$ might be sufficient. Also, in the coded case, soft information should be provided to the decoder and this further complicates the decoder. Therefore, using rotations of dimension smaller than B, may yield a desirable tradeoff between diversity, rate, and complexity.

Manuscript received May 24, 2007; revised October 9, 2007. The work by A. Guillén i Fàbregas was supported in part by the Australian Research Council under ARC Grants DP0558861 and RN0459498. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Nice, France, June 2007.

A. Guillén i Fàbregas is with the Signal Processing and Communications Laboratory, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, U.K. (e-mail: guillen@ieee.org).

G. Caire is with the Electrical Engineering Department, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: caire@usc.edu).

Communicated by H. Boche, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2008.920244

In this correspondence, we study the *reliability exponent*, namely, the optimal exponent of the error probability of such schemes with the signal-to-noise ratio (SNR) in a logarithmic scale, and illustrate the rate–diversity–complexity tradeoff for coded modulation schemes constructed over multidimensional signal sets.

II. System Model

We consider a single-input single-output block-fading channel with B fading blocks, whose system model is given by the following:

$$\boldsymbol{y}_b = \sqrt{\rho} h_b \boldsymbol{x}_b + \boldsymbol{z}_b, \qquad b = 1, \dots, B \tag{1}$$

where $h_b \in \mathbb{C}$ is the bth fading coefficient, $\boldsymbol{y}_b \in \mathbb{C}^L$ is the received signal vector corresponding to fading coefficient b, $\boldsymbol{x}_b \in \mathbb{C}^L$ is the portion of codeword allocated to block b, and $\boldsymbol{z}_b \in \mathbb{C}^L$ is the vector of independent and identically distributed (i.i.d.) noise samples $\sim \mathcal{N}_{\mathbb{C}}(0,1)$. The transmitted signal is normalized in energy, i.e., $\mathbb{E}[|x|^2] = 1$. Hence, ρ is the average received SNR. We assume that the fading coefficients are i.i.d. from block to block and from codeword to codeword, and that they are perfectly known at the receiver, with Nakagami-m-distributed coefficients, i.e.,

$$p_{|h|}(\xi) = 2m^m \xi^{2m-1} \Big(\Gamma(m)\Big)^{-1} e^{-m\xi^2}$$

for $m > 0$ where $\Gamma(\xi) \triangleq \int_0^{+\infty} t^{\xi-1} e^{-t} dt$ is the Gamma function. Nakagami-m fading characterizes a large class of statistics, including Rayleigh, by setting $m = 1$, and Rician with parameter $\mathcal{K}$ by setting $m = (\mathcal{K}+1)^2/(2\mathcal{K}+1)$. We define $\gamma_b \triangleq |h_b|^2$, $b = 1, \dots, B$. We can express (1) in matrix form as

$$\boldsymbol{Y} = \sqrt{\rho} \boldsymbol{H} \boldsymbol{X} + \boldsymbol{Z} \tag{2}$$

where

$$\boldsymbol{Y} = [\boldsymbol{y}_1, \dots, \boldsymbol{y}_B]^T \in \mathbb{C}^{B \times L}$$
$$\boldsymbol{X} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_B]^T = [\boldsymbol{X}_1, \dots, \boldsymbol{X}_L] \in \mathbb{C}^{B \times L}$$
$$\boldsymbol{Z} = [\boldsymbol{z}_1, \dots, \boldsymbol{z}_B]^T \in \mathbb{C}^{B \times L}$$

and

$$\boldsymbol{H} = \mathrm{diag}(h_1, \dots, h_B) \in \mathbb{C}^{B \times B}.$$

Codewords $\boldsymbol{X}$ form a coded modulation scheme $\mathcal{X} \subset \mathbb{C}^{B \times L}$. We consider that $\mathcal{X}$ is obtained as the concatenation of a binary code $\mathcal{C} \in \mathbb{F}_2^n$ of rate r, a modulation over the signal constellation $\mathcal{S} \in \mathbb{C}$ with $M = \log_2 |\mathcal{S}|$, and K rotations $\boldsymbol{M}_k \in \mathbb{C}^{N \times N}$ with $KN = B$. In particular, we have that

$$\boldsymbol{x}_{\ell,k} = \boldsymbol{M}_k \boldsymbol{s}_{\ell,k}, \qquad \ell = 1, \dots, L \tag{3}$$

where $\boldsymbol{s}_{\ell,k} = (s_{\ell,k,1}, \dots, s_{\ell,k,N})^T \in \mathcal{S}^N$ is the vector of signal constellation symbols that is rotated by the kth rotation matrix, $\boldsymbol{x}_{\ell,k} = (x_{\ell,k,1}, \dots, x_{\ell,k,N})^T$ is the portion of transmitted signal at the ℓth channel use that has been rotated by the kth rotation, and $\boldsymbol{x}_\ell = \left[\boldsymbol{x}_{\ell,1}^T, \dots, \boldsymbol{x}_{\ell,K}^T\right]^T$ is the transmitted signal at the ℓth channel use. The rotation matrices are constrained to be unitary, i.e., $\boldsymbol{M}_k \boldsymbol{M}_k^\dagger = \boldsymbol{I}$. We will be interested in *full-diversity* rotations, i.e., rotation matrices $\boldsymbol{M}$ for which

$$\boldsymbol{M}(\boldsymbol{s} - \boldsymbol{s}') \neq \boldsymbol{0}, \qquad \forall\, \boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}^N,\ \boldsymbol{s} \neq \boldsymbol{s}' \tag{4}$$

componentwise. This implies that, if the vector $\boldsymbol{s} - \boldsymbol{s}'$ has any number of nonzero components, its rotated version $\boldsymbol{M}(\boldsymbol{s} - \boldsymbol{s}')$ will have *all* nonzero components. Reference [4] reports rotation matrices using the row convention while here we use a column convention.

The rate in bits per channel use of this scheme is independent of N, and is given by $R = rM$. This general formulation includes the case where only one single rotation of dimension B is used, as well as the other extreme, with B trivial rotations of dimension $N = 1$ (the nonrotated case). As we shall see in the following, while the rate is independent of N, the reliability exponent will depend on N.

Definition 1: The block diversity of a coded modulation scheme $\mathcal{X} \subset \mathbb{C}^{B\times L}$ is defined as

$$\delta = \min_{\boldsymbol{X}(i),\boldsymbol{X}(j)\in\mathcal{X},j\neq i} |\{b \in (1,\ldots,B)|\boldsymbol{x}_b(i) \neq \boldsymbol{x}_b(j)\}|. \tag{5}$$

In words, the block diversity is the minimum number of nonzero rows of $\boldsymbol{X}(i) - \boldsymbol{X}(j)$ for any pair of codewords $\boldsymbol{X}(j) \neq \boldsymbol{X}(i) \in \mathcal{X}$.

Proposition 1: Given a coded modulation scheme $\mathcal{X} \subset \mathbb{C}^{B\times L}$, the block diversity is upper-bounded by

$$\delta \leq N\left(1 + \left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right). \tag{6}$$

Proof: The result follows from the straightforward application of the Singleton bound to the coded modulation $\mathcal{X}$ seen as a code of block length K, over an alphabet of size 2^{MNL}. □

We will say that a code is blockwise maximum-distance separable (MDS) if it attains the Singleton bound of Proposition 1 with equality.

III. Outage Probability

The channel defined in (1) is not information stable and has zero capacity for any finite B [5], since there is a nonzero probability that the transmitted message is detected in error. For sufficiently large L, the word error probability $P_e(\rho, \mathcal{X})$ of any coding scheme $\mathcal{X} \subset \mathbb{C}^{B\times L}$ is lowe-bounded by the information outage probability [2]

$$P_e(\rho, \mathcal{X}) \geq P_{\text{out}}(\rho, R) \triangleq \Pr(I(\rho, \boldsymbol{H}) \leq R) \tag{7}$$

where $I(\rho, \boldsymbol{H})$ is the input–output mutual information of the channel for a given fading realization $\boldsymbol{H}$ (in bits per channel use). In this work, we will study the behavior of $P_{\text{out}}(\rho, R)$ for large ρ, for which the optimal power allocation when no channel state information (CSI) is available at the transmitter, corresponds to evenly distributing the available power across all B blocks. In the case of uniform allocation, and for a fixed $\boldsymbol{H}$, the outage probability is minimized when the entries of $\boldsymbol{X} \in \mathcal{X}$ are i.i.d. Gaussian $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$. In this case

$$I(\rho, \boldsymbol{H}) = \frac{1}{B}\sum_{b=1}^{B} \log_2(1 + \rho\gamma_b).$$

When the scheme described in Section II is used (assuming uniform inputs), we can write that

$$I(\rho, \boldsymbol{H}) = \frac{1}{K}\sum_{k=1}^{K} \frac{1}{N} I_k(\rho, \boldsymbol{H}_k) = \frac{1}{B}\sum_{k=1}^{K} I_k(\rho, \boldsymbol{H}_k)$$

where the mutual information of the $N \times N$ MIMO channel induced by the kth rotation is (see, e.g., [6])

$$I_k(\rho, \boldsymbol{H}_k) = MN - \frac{1}{2^{MN}} \sum_{\boldsymbol{s}\in\mathcal{S}^N} \mathbb{E}_{\boldsymbol{z}}\left[\log_2\left(\sum_{\boldsymbol{s}'\in\mathcal{S}^N} e^{-\|\sqrt{\rho}\boldsymbol{H}_k\boldsymbol{M}_k(\boldsymbol{s}-\boldsymbol{s}')+\boldsymbol{z}\|^2+\|\boldsymbol{z}\|^2}\right)\right] \tag{8}$$

and $\boldsymbol{H}_k = \operatorname{diag}(h_{(k-1)N+1}, \ldots, h_{kN}) \in \mathbb{C}^{N\times N}$ are the channel coefficients *used* by rotation k, and $\boldsymbol{z} \in \mathbb{C}^N$ is a dummy additive white Gaussian noise (AWGN) vector over which the expectation is computed.

Fig. 1 shows the mutual information with Gaussian inputs, unrotated 16-QAM, and rotated 16-QAM in a block-fading channel with $B = 4$ blocks and $h_1 = 1.5$ and $h_2 = h_3 = h_4 = 0.1$. This choice of the channel coefficients is particularly interesting since three

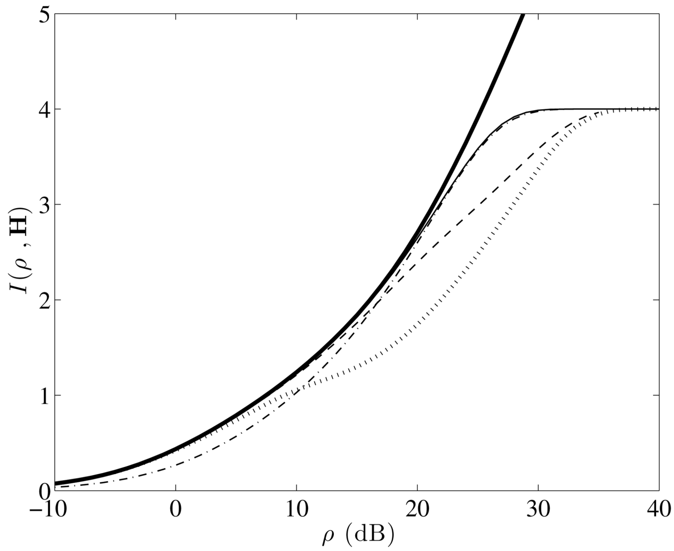

Fig. 1. Instantaneous mutual information $I(\rho, \boldsymbol{H})$ (bits/channel use) in a block-fading channel with $B = 4$ blocks and $h_1 = 1.5$ and $h_2 = h_3 = h_4 = 0.1$ with Gaussian inputs (thick solid) and rotated 16-QAM inputs with the optimal Krüskemper (thin solid), mixed (thin dash–dotted), two independent two-dimensional cyclotomic rotations (thin dashed) and no rotations (thick dotted).

out of the four components are in a deep fade.[1] Rotations of dimension N yield vanishing (for large ρ) error probability when there are up to $N - 1$ deeply faded blocks [1]. The mutual information achieved by the rotated 16-QAM is very close to that attained by the Gaussian distribution in a wider range of ρ than uncoded 16-QAM. At ρ 25 dB , the Krüskemper rotation gains 1 bit of information with respect to unrotated 16-QAM. Combining two cyclotomic rotations of dimension $N = 2$ brings also significant information gains with respect to unrotated 16-QAM. As we shall see, this effect brings substantial exponent benefits with respect to the unrotated case. We also observe a difference between Krüskemper and mixed (2×2) rotations, especially at low rates, while unrotated 16-QAM performs almost as well with much less decoding complexity. Rotations provide only mutual information advantages at high rates.

IV. Optimal Reliability

We define the diversity reliability exponent of a given coded modulation scheme $\mathcal{X}$ as

$$d_{\mathcal{X}} = \lim_{\rho\to+\infty} -\frac{\log P_e(\rho, \mathcal{X})}{\log \rho} \tag{9}$$

and the optimal diversity reliability exponent is $d^\star \triangleq \sup_{\mathcal{X}} d_{\mathcal{X}}$.

When no particular structure is imposed on the coded modulation scheme $\mathcal{X}$, we have the following result.

Lemma 1: The diversity reliability exponent of any scheme $\mathcal{X}$ subject to the constraint $\frac{1}{BL}\mathbb{E}[\|\boldsymbol{X}\|^2] \leq 1$ is upper-bounded by $d_{\mathcal{X}} \leq d^\star = mB$. The optimal diversity reliability exponent is achieved by random Gaussian codes of rate $R > 0$ with entries $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$. The optimal exponent $d^\star$ can also be achieved by random coded modulation schemes $\mathcal{X}$ of rate R consisting of a random coded modulation scheme over a complex-plane set $\mathcal{S}$ of size $|\mathcal{S}| = 2^M$ concatenated with a full-diversity rotation of dimension B, whenever $0 \leq \frac{R}{M} < 1$.

Proof: The converse is proved in [7], [8]. Furthermore, [7], [8] also show that the random Gaussian ensemble achieves the optimal

[1]In a nonergodic scenario, the ergodic information rate averaged over the channel realizations does not have a practical relevance. Instead, the *bad* channels dominate the outage probability at large ρ.

exponent. Appendix II, by letting $N = B$, shows the proof that the random coded modulation scheme over a single full-diversity rotation of dimension B achieves the same exponent. □

We have included the achievability with the random coded modulation ensemble over the B-dimensional rotated constellation to illustrate that a coding scheme with discrete inputs can also achieve the optimal exponent. This result which is based on a divide and conquer approach, should be rather intuitive: the rotation of dimension B achieves full diversity while the coding gain is then left to the outer coded modulation scheme over $\mathcal{S}$. With no rotations, the optimal diversity reliability exponent is given by the Singleton bound [8]

$$d^\star = m\left(1 + \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor\right). \quad (10)$$

The advantage of rotations is clear: they can achieve the optimal diversity reliability exponent for the whole range of rates. With no rotations, the largest rate such that the optimal exponent is achieved is $R = \frac{M}{B}$. The following result characterizes the optimal exponent when rotations of smaller size $N < B$ are employed.

Proposition 2: The diversity reliability exponent for the coded modulation schemes based on K rotations of dimension N, in a Nakagami-m block-fading channel with $B = KN$ blocks is upper-bounded by

$$d_{\mathcal{X}} \le mN\left(1 + \left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right). \quad (11)$$

Proof: See Appendix I. □

Proposition 3: The diversity reliability exponent in a Nakagami-m block-fading channel with $B = KN$ of random coded modulation schemes based on K rotations of dimension N of length L satisfying $\lim_{\rho\to\infty} \frac{L}{\rho} = \lambda$, is lower-bounded by $d_{\mathcal{X}} \ge \lambda BM \log 2\left(1 - \frac{R}{M}\right)$ when $0 \le \lambda NM \log 2 < m$ and otherwise by

$$d_{\mathcal{X}} \ge \min\left\{ mN\left\lceil \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rceil, mN\left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor + \lambda M \log 2\left(B\left(1 - \frac{R}{M}\right) - N\left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right)\right\}. \quad (12)$$

Proof: See Appendix II. □

The proof of the last two propositions closely follows the reasoning of [7], [8]. The preceding results lead to the following theorem.

Theorem 1: The optimal diversity reliability exponent for the coded modulation schemes based on K rotations of dimension N, in a Nakagami-m block-fading channel with $B = KN$ blocks is

$$d^\star_{\mathcal{X}} = mN\left(1 + \left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right) \quad (13)$$

whenever $\frac{B}{N}\left(1 - \frac{R}{M}\right)$ is not an integer.

Proof: Proposition 2 shows that

$$d_{\mathcal{X}} \le mN\left(1 + \left\lfloor \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right).$$

Letting $\lambda \to \infty$ in Proposition 1 shows that $d_{\mathcal{X}} \ge mN\left\lceil \frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rceil$. Noting that $\lceil x \rceil = \lfloor x \rfloor + 1$ whenever x is not an integer leads the desired result. □

Theorem 1 gives a dual result to that of [8] and shows that the optimal exponent is given by m times the Singleton bound of (6), proving its optimality and separating the roles of the channel distribution and of the code construction. The optimal codes are *blockwise* MDS. For $N > 1$, Theorem 1 suggests that the optimal coding scheme is to use an MDS coded modulation scheme constructed over $\mathcal{S}$ in a block-fading channel with $K = \frac{B}{N}$ blocks, concatenated with K rotations of dimension N. In this case, the MDS constraint on the code is relaxed, since it has to be MDS for a smaller number of blocks. Theorem 1 implicitly introduces an equivalent channel model, namely, a block-fading channel with $K = \frac{B}{N}$, where each block has diversity mN. When $K = 1$, $N = B$, there is only one single rotation of *full* dimension, Theorem 1 generalizes Lemma 1. Theorem 1 generalizes and proves the optimality of the modified Singleton bound introduced in [12].

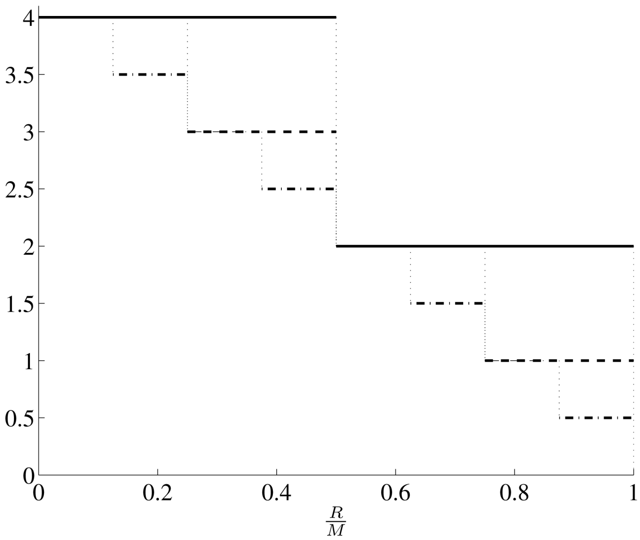

Fig. 2. Reliability exponents for $B = 8, m = 0.5$, and rotations of dimensions $N = 1$ (dash-dotted), $N = 2$ (dashed), and $N = 4$ (solid).

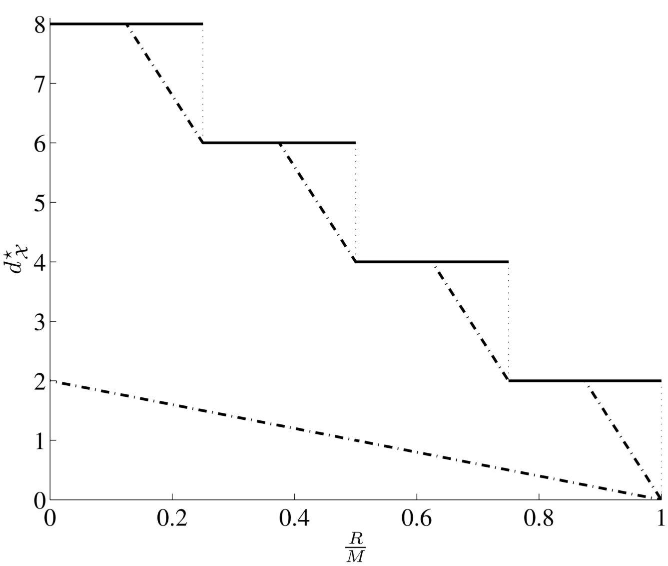

Fig. 3. Reliability exponents for $B = 8, m = 1$, and rotations of dimensions $N = 2$. The random coding exponents for $\lambda M \log 2 = \frac{m}{2N}$ (lower dash–dotted curve) and $\lambda M \log 2 = \frac{4m}{N}$ (upper dash–dotted curve) are also shown.

Fig. 2 shows the exponents for $B = 8$, $m = 0.5$, and $N = 1, 2, 4$. The figure confirms the intuition that rotations should increase the exponent. For example, for $\frac{R}{M} = \frac{1}{2}$, we have that unrotated inputs yield $d^\star_{\mathcal{X}} = m5$, while for rotations with $N = 2$, then $d^\star_{\mathcal{X}} = m6$ and for $N = 4$ $d^\star_{\mathcal{X}} = m8$, i.e., full diversity. The task of achieving diversity is hence split between the code $\mathcal{C}$ and the rotations. Fig. 3 shows the upper bound as well as the random coding lower bound given in Propositions 2 and 3, respectively. As we see, if λ is increased, both bounds coincide

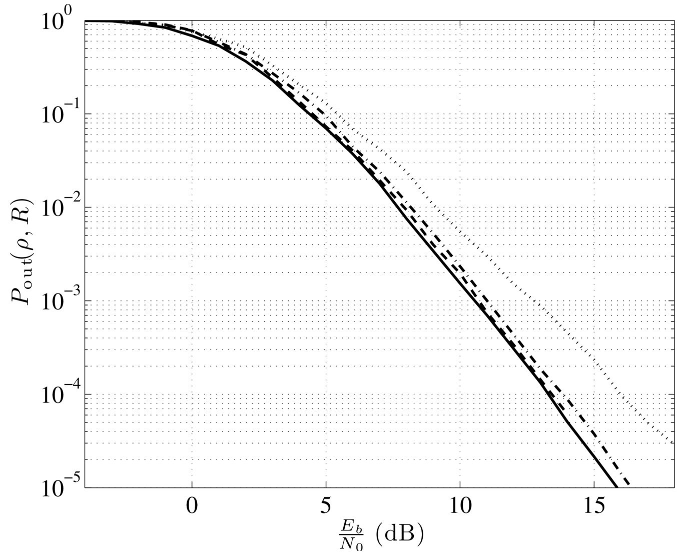

Fig. 4. Outage probability for $R = 1$ bits per channel use in a block-fading channel with $B = 4$, $m = 1$, with Gaussian (solid line), rotated quaternary phase-shift keying (QPSK) inputs with one Krüskemper rotation of dimension $N = 4$ (dashed line), rotated QPSK inputs with two cyclotomic rotations of dimension $N = 2$ (dash–dotted) and unrotated QPSK inputs (dotted).

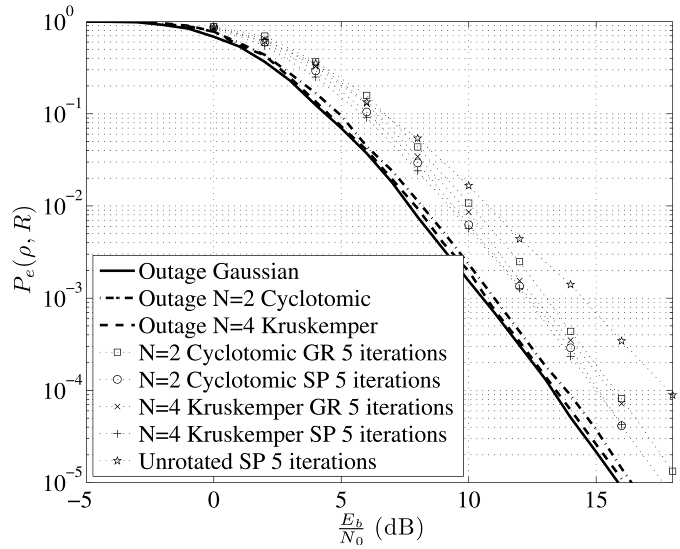

Fig. 6. Error probability for $R = 1$ bits per channel use in a block-fading channel with $B = 4$, $m = 1$ using the $(5, 7)_8$ convolutional code, and QPSK modulation with Gray (GR) and set-partitioning (SP) labeling. The outage probabilities with Gaussian inputs (thick solid line), rotated QPSK with one Krüskemper rotation of dimension $N = 4$ (dashed line), rotated QPSK with two cyclotomic rotations of dimension $N = 2$ (dash–dotted) are also shown.

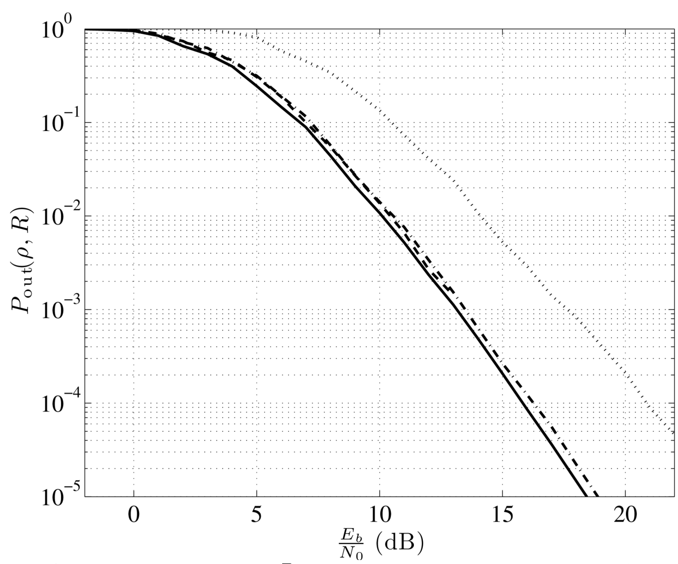

Fig. 5. Outage probability for $R = 2$ bits per channel use in a block-fading channel with $B = 4$, $m = 1$, with Gaussian (solid line), rotated 16-QAM inputs with one Krüskemper rotation of dimension $N = 4$ (dashed line), rotated 16-QAM inputs with two cyclotomic rotations of dimension $N = 2$ (dash–dotted), and unrotated 16-QAM inputs (dotted).

in a larger support. Eventually, for $\lambda \to \infty$ they coincide where they are continuous.

To illustrate the benefits of rotations, Figs. 4 and 5 show $P_{\rm out}(\rho, R)$ as a function of $\frac{E_b}{N_0}$ for $m = 1$ and $B = 4$ for $R = 2$. Gaussian inputs achieve the optimal exponent, namely, $d^\star = B = 4$, while unrotated inputs have $d^\star_{\mathcal{X}} = 3$ [7]. As we observe, using two rotations of dimension $N = 2$, not only allows to recover the largest possible exponent (in agreement with Theorem 1) but also brings a large gain. Using a rotation of dimension $N = 4$ incurs much larger complexity and does not bring any exponent nor gain improvements.

To illustrate that the above theoretical results are approachable with practical schemes, Fig. 6 shows the error probability of rotated and unrotated quadrature phase-shift keying (QPSK) modulation using the $(5, 7)_8$ convolutional code with 128 information bits per frame. In the case of two rotations of dimension $N = 2$, we separately use bit-interleaved coded modulation (BICM) [9] followed by a rotation on the outputs generated by generator polynoimial 5_8 and 7_8. Since the $(5, 7)_8$ convolutional code has full diversity in a block-fading channel with $K = 2$ blocks, the overall scheme has full diversity. A similar construction can be obtained using blockwise concatenated codes [7] or multiplexed turbo codes [10]. These schemes will closely approach the outage probability of the channel for any (sufficiently large) block length. Rotated systems compute the metrics for all the candidate points [6]. Again, here the gain obtained by rotations is significant. All rotated systems show a steeper slope to that of the unrotated case. Moreover, we observe that using a rotation of full dimension $N = 4$ yields once more a small gain with respect to using two rotations of dimension $N = 2$, while significantly increasing the decoding complexity. We also see that, set-partitioning labeling yields some performance advantage over Gray. From results not shown here, both Gray and set partitioning show improved performance with the iterations. This is due to the fact that rotations induce an equivalent MIMO channel, and the iterative decoder assists in iteratively removing the self-interference introduced by the rotation.

V. Conclusion

We have studied coded modulation schemes over Nakagami-m block-fading channels with discrete constellations. We have derived the optimal exponent for multidimensional constellations obtained from the rotation of complex-plane constellations, and we have shown that there is a tradeoff between the transmission rate, optimal diversity, dimension of the rotations, and size of the complex-plane constellation, given by a modified form of the Singleton bound. Since rotations induce an increased decoding complexity, the Singleton bound establishes the optimal rate–diversiy–complexity tradeoff. Practical coding schemes are shown to achieve the optimal tradeoff.

Appendix I
Proof of Proposition 2

We write $f(z) \doteq z^d$ to indicate that $\lim_{z\to\infty} \frac{\log f(z)}{\log z} = d$. The exponential inequalities $\dot{\geq}$ and $\dot{\leq}$ are defined similarly [11]. For vec-

$$\overline{P_e(\rho)} \dot{\leq} \int_{\boldsymbol{\alpha}\in\mathcal{E}_\delta\cap\mathbb{R}_+^B} \rho^{-m\sum_{b=1}^B \alpha_b} d\boldsymbol{\alpha} \tag{20}$$

$$+ \int_{\boldsymbol{\alpha}\in\mathcal{E}_\delta^c\cap\mathbb{R}_+^B} \exp\left(-\log\rho\left[m\sum_{b=1}^B \alpha_b + \lambda BM\log 2 E_\delta(\boldsymbol{\alpha})\right]\right) d\boldsymbol{\alpha}. \tag{21}$$

tors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, the notation $\boldsymbol{x} \prec \boldsymbol{y}$ denotes componentwise inequality $x_i < y_i, i = 1, \dots, n$. The inequalities $\succ, \preceq, \succeq$ are used similarly. The function $\mathbb{1}\{\mathcal{E}\}$ is the indicator function of the event $\mathcal{E}$. The normalized fading coefficients are defined as $\alpha_b \triangleq -\frac{\log\gamma_b}{\log\rho}$ for $b = 1, \dots, B$ [11]. From [8] we have that the joint distribution of the vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_B)$ in the limit for large ρ behaves as $p(\boldsymbol{\alpha}) \doteq \rho^{-m\sum_{b=1}^B \alpha_b}$ for $\boldsymbol{\alpha} \in \mathbb{R}_+^B$. The kth vector of normalized fading coefficients is defined as $\boldsymbol{\alpha}_k \triangleq (\alpha_{N(k-1)+1}, \dots, \alpha_{Nk})$.

Since rotations induce an MIMO channel, we bound (8) as

$$I(\rho, \boldsymbol{H}) \leq \frac{1}{K}\sum_{k=1}^K \frac{1}{N}\min\left\{NM, \log\det(\boldsymbol{I} + \rho\boldsymbol{H}_k\boldsymbol{M}_k\boldsymbol{M}_k^\dagger\boldsymbol{H}_k^\dagger)\right\}$$
$$= \frac{1}{K}\sum_{k=1}^K \min\left\{M, \frac{1}{N}\sum_{n=1}^N \log(1+\rho\gamma_{N(k-1)+n})\right\}.$$

Now, we can express the outage probability for large ρ as

$$P_{\text{out}}(\rho, R)$$
$$\dot{\geq} \Pr\left(\frac{1}{K}\sum_{k=1}^K \min\left\{M, \frac{\log\rho}{N}\sum_{n=1}^N [1-\alpha_{N(k-1)+n}]_+\right\} < R\right)$$
$$\dot{\geq} \int_{\mathcal{O}_\epsilon\cap\mathbb{R}_+^B} \rho^{-m\sum_{b=1}^B \alpha_b} d\boldsymbol{\alpha} \tag{14}$$

where the first exponential inequality follows from

$$(1+\rho\gamma_{N(k-1)+n}) \doteq [1-\alpha_{N(k-1)+n}]_+$$

where $[x]_+ = \max(0, x)$ denotes the positive part of $x \in \mathbb{R}$,

$$\mathcal{O}_\epsilon \triangleq \left\{\boldsymbol{\alpha} \in \mathbb{R}^B : \frac{1}{K}\sum_{k=1}^K \mathbb{1}\{\boldsymbol{\alpha}_k \succeq \boldsymbol{1}+\boldsymbol{\epsilon}\} > 1 - \frac{R}{M}\right\}$$

denotes the large ρ outage event, and where $\boldsymbol{1} = (1, \dots, 1)$ and $\boldsymbol{\epsilon} = (\epsilon, \dots, \epsilon)$ both of dimension N. Equation (2) is valid for any $\epsilon > 0$ and in particular for $\epsilon \to 0$. Using Varadhan's lemma, we have

$$d_{\mathcal{X}} \leq d_{\text{out}} = \inf_{\mathcal{O}_\epsilon\cap\mathbb{R}_+^B}\left\{m\sum_{b=1}^B \alpha_b\right\}. \tag{15}$$

It is not difficult to show that $d_{\text{out}} = m\kappa N$, where κ is the unique integer such that $\kappa < K\left(1-\frac{R}{M}\right) \leq \kappa + 1$. This yields the result.

Appendix II
Proof of Proposition 1

For any two codewords $\boldsymbol{X}(0), \boldsymbol{X}(1) \in \mathcal{X}$, we can write that the pairwise error probability

$$P(\boldsymbol{X}(0)\to\boldsymbol{X}(1)|\boldsymbol{H}) \leq \prod_{k=1}^K \exp\left(-\frac{\rho}{4}\left\|\boldsymbol{H}_k\boldsymbol{M}_k(\boldsymbol{S}_k(0)-\boldsymbol{S}_k(1))\right\|^2\right)$$

where $\boldsymbol{S}_k(i)$ is such that the portion of codeword rotated by the kth matrix is $\boldsymbol{X}_k(i) = \boldsymbol{M}_k\boldsymbol{S}_k(i)$, and $\boldsymbol{H} = \text{diag}(\boldsymbol{H}_1, \dots, \boldsymbol{H}_K)$. Assuming uniform i.i.d. entries of $\boldsymbol{S}_k(0)$ and $\boldsymbol{S}_k(1)$, the ensemble pairwise error probability can be expressed as

$$\overline{P(\boldsymbol{X}(0)\to\boldsymbol{X}(1)|\boldsymbol{H})}$$
$$\leq \prod_{k=1}^K\left[\frac{1}{2^{2MN}}\sum_{\boldsymbol{s}\in\mathcal{S}^N}\sum_{\boldsymbol{s}'\in\mathcal{S}^N}\exp\left(-\frac{\rho}{4}\left\|\boldsymbol{H}_k\boldsymbol{M}_k(\boldsymbol{s}-\boldsymbol{s}')\right\|^2\right)\right]^L. \tag{16}$$

Similarly to [7], summing over the $2^{LBR}-1$ codewords different from the 0 message we have that

$$\overline{P_e(\rho|\boldsymbol{H})} = \exp\left(-BLM\log 2E(\rho, \boldsymbol{\alpha})\right) \tag{17}$$

where the exponent $E(\rho, \boldsymbol{\alpha})$ is given by

$$E(\rho, \boldsymbol{\alpha}) = 1 - \frac{R}{M}$$
$$-\frac{1}{BM}\sum_{k=1}^K \log_2\left(1+\frac{1}{2^{MN}}\sum_{\boldsymbol{s}'\neq\boldsymbol{s}} e^{-\frac{1}{4}\sum_{n=1}^N \rho^{1-\alpha_{N(k-1)+n}}|\tilde{x}_{k,n}|^2}\right)$$

and $\tilde{\boldsymbol{x}}_k = \boldsymbol{M}_k(\boldsymbol{s}-\boldsymbol{s}') = (\tilde{x}_{k,1}, \dots, \tilde{x}_{k,N})^T$ is the rotated difference vector. We now assume that the rotation matrices have *full diversity*. That implies that all the components of the rotated difference vector $\tilde{\boldsymbol{x}}_k$ are different from zero. For large ρ, we have that

$$\lim_{\rho\to\infty}\log_2\left(1+\frac{1}{2^{MN}}\sum_{\boldsymbol{s}'\neq\boldsymbol{s}} e^{-\frac{1}{4}\sum_{n=1}^N \rho^{1-\alpha_{N(k-1)+n}}|\tilde{x}_{k,n}|^2}\right) \tag{18}$$

is equal to MN if $\boldsymbol{\alpha}_k \succ \boldsymbol{1}$ and 0 otherwise, where $\boldsymbol{\alpha}_k = (\alpha_{N(k-1)+1}, \dots, \alpha_{Nk})^T$. Hence

$$\overline{P_e(\rho|\boldsymbol{H})} \dot{\leq} e^{-BLM\log 2E_\delta(\boldsymbol{\alpha})} \tag{19}$$

where

$$E_\delta(\boldsymbol{\alpha}) = 1 - \frac{R}{M} - \frac{1}{K}\sum_{k=1}^K \mathbb{1}\{\boldsymbol{\alpha}_k \succeq \boldsymbol{1}-\boldsymbol{\delta}\}$$

and $\boldsymbol{\delta} = (\delta, \dots, \delta) \in \mathbb{R}_+^N$. We now define the large ρ error event as $\mathcal{E}_\delta = \left\{\boldsymbol{\alpha} \in \mathbb{R}^B : E_\delta(\boldsymbol{\alpha}) \leq 0\right\}$. Using the previous results we write that

$$\overline{P_e(\rho)} \dot{\leq} \int_{\boldsymbol{\alpha}\in\mathbb{R}_+^B} \rho^{-m\sum_{b=1}^B \alpha_b}\min\left\{1, e^{-BLM\log 2E_\delta(\boldsymbol{\alpha})}\right\} d\boldsymbol{\alpha}$$
$$= \int_{\boldsymbol{\alpha}\in\mathcal{E}_\delta\cap\mathbb{R}_+^B} \rho^{-m\sum_{b=1}^B \alpha_b} d\boldsymbol{\alpha}$$
$$+ \int_{\boldsymbol{\alpha}\in\mathcal{E}_\delta^c\cap\mathbb{R}_+^B} \rho^{-m\sum_{b=1}^B \alpha_b} e^{-BLM\log 2E_\delta(\boldsymbol{\alpha})} d\boldsymbol{\alpha}.$$

In a similar way to the proof of Lemma 1 the probability of two randomly chosen codewords over $\mathcal{S}$ being the same is strictly greater than zero, and goes to zero only for $L \to \infty$. We now study how large L has to be in order for this event not to dominate the overall error probability. If we let $\lambda = \lim_{\rho\to\infty}\frac{L}{\log\rho}$ we can write (20)–(21) at the top of the page. Therefore, the overall random coding exponent is given by the minimum of the exponents of (20) and (21)

$$d_{\mathcal{X}}(R) \geq d_{\mathcal{X}}^{(r)}(R) = \sup_{\delta>0}\min\left\{d_{\mathcal{X}}^{(r),\infty}(R), d_{\mathcal{X}}^{(r),\lambda}(R)\right\} \tag{22}$$

$$d_{\mathcal{X}}^{(r),\lambda}(R) = \lambda BM \log 2\left(1-\frac{R}{M}\right) + \inf_{\boldsymbol{\alpha}\in\mathcal{E}_\delta^c\cap\mathbb{R}_+^B}\left\{m\sum_{b=1}^{B}\alpha_b - \lambda BM\log 2\frac{1}{K}\sum_{k=1}^{K}\mathbb{1}\{\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}\}\right\}$$

$$= \lambda BM \log 2\left(1-\frac{R}{M}\right) + m\inf_{\boldsymbol{\alpha}\in\mathcal{E}_\delta^c\cap\mathbb{R}_+^B}\left\{\sum_{k=1}^{K}\left(\sum_{n=1}^{N}\alpha_{k,n} - \frac{\lambda NM\log 2}{m}\mathbb{1}\{\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}\}\right)\right\}$$

where

$$d_{\mathcal{X}}^{(r),\infty}(R) = \inf_{\boldsymbol{\alpha}\in\mathcal{E}_\delta\cap\mathbb{R}_+^B} m\left\{\sum_{b=1}^{B}\alpha_b\right\}$$

is the exponent corresponding to (20) and

$$d_{\mathcal{X}}^{(r),\lambda}(R) = \inf_{\boldsymbol{\alpha}\in\mathcal{E}_\delta^c\cap\mathbb{R}_+^B}\left\{m\sum_{b=1}^{B}\alpha_b + \lambda BM\log 2\ E_\delta(\boldsymbol{\alpha})\right\}$$

is the exponent that characterizes the effect of finite length (21). It is not difficult to show that the first infimum is achieved by κ vectors $\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}$, where κ is the unique integer such that $\kappa-1 < \left\lceil K\left(1-\frac{R}{M}\right)\right\rceil \le \kappa$ resulting in the exponent being

$$d_{\mathcal{X}}^{(r),\infty}(R) = (1-\delta)mN\left\lceil\frac{B}{N}\left(1-\frac{R}{M}\right)\right\rceil. \quad (23)$$

As for the second exponent, we can rewrite it as the expression shown at the top of the page, where

$$\mathcal{E}_\delta^c \triangleq \left\{\boldsymbol{\alpha}\in\mathbb{R}^B : \sum_{k=1}^{K}\mathbb{1}\{\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}\} < K\left(1-\frac{R}{M}\right)\right\}.$$

We distinguish two cases. When $0 \le \lambda NM\log 2 < m$ then

$$\sum_{n=1}^{N}\alpha_{k,n} - \frac{\lambda NM\log 2}{m}\mathbb{1}\{\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}\} \quad (24)$$

attains its minimum value for $\boldsymbol{\alpha}_k = \mathbf{0}$. When $\lambda NM\log 2 \ge m$, the constraint set dictates that there should be $\kappa = \left\lfloor K\left(1-\frac{R}{M}\right)\right\rfloor$ vectors $\boldsymbol{\alpha}_k \succeq \mathbf{1}-\boldsymbol{\delta}$, and the infimum becomes

$$\lambda BM\log 2\left(1-\frac{R}{M}\right) + m\left\lfloor K\left(1-\frac{R}{M}\right)\right\rfloor\left(N(1-\delta) - \frac{\lambda NM\log 2}{m}\right). \quad (25)$$

Combining the previous results and noting that the supremum in (22) is achieved for $\delta \to 0$, we find the desired result.

References

[1] F. Oggier and E. Viterbo, "Algebraic number theory and code design for rayleigh fading channels," *Foundations and Trends in Communications and Information Theory*, vol. 1, pp. 333–415, 2004.

[2] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[3] E. Viterbo and J. Boutros, "A universal lattice decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1639–1642, Jul. 1999.

[4] E. Viterbo and F. Oggier, "Tables of Algebraic Rotations," [Online]. Available: http://www.tlc.polito.it/~ viterbo

[5] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

[6] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[7] A. Guillén i Fàbregas and G. Caire, "Coded modulation in the block-fading channel: Coding theorems and code construction," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 262–271, Jan. 2006.

[8] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, "A tight lower bound to the outage probability of block-fading channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4314–4322, Nov. 2007.

[9] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[10] J. J. Boutros, E. C. Strinati, and A. Guillén i Fàbregas, "Turbo code design for block fading channels," in *Proc. 42nd Annu. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep./Oct. 2004.

[11] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[12] N. Gresset, J. J. Boutros, and L. Brunel, "Optimal linear precoding for BICM over MIMO channels," in *Proc IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 66.

The Average Value for the Probability of an Undetected Error

Patrick N. Perry, *Member, IEEE*, and
Marc P. C. Fossorier, *Fellow, IEEE*

***Abstract*—A new identity for the weight enumerator of a code is derived. The new identity is shown to be related to the average value of the probability of an undetected error when p is a continuous random variable.**

***Index Terms*—Linear codes, probability of an undetected error, weight enumerators.**

I. Introduction

The weight enumerator for a length-n linear code is the polynomial $A(x) = A_0 + A_1x + A_2x^2 + \cdots + A_nx^n$ where A_i is the number of codewords of weight i in the code. We derive a new identity for the sum of the code weight enumerator coefficients A_i scaled by the binomial coefficients $\binom{n}{i}$. The resulting identity is in terms of the dual code weight enumerator coefficients $A_i^\perp$.

The probability of an undetected error for a linear code when using a binary-symmetric channel (BSC) with a probability of an error being p is known to be $P_{\mathrm{ue}}(C,p) = \sum_{i=1}^{n} A_i\, p^i\,(1-p)^{n-i}$, see [1], [2].

Manuscript received April 14, 2007; revised December 8, 2007. This work was supported by the Hawaii Pacific University Trustees' Scholarly Endeavors Program. The material in this correspondence was presented in part at the 2006 IEEE Information theory Workshop, Chengdu, China, October 2006.

P. N. Perry is with the Mathematics Department, Hawaii Pacific University, Honolulu, HI USA (e-mail: pperry@hpu.edu).

M. P. C. Fossorier was with the Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822 USA (e-mail: mfossorier2 @yahoo.com).

Communicated by I. Dumer, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.920241