TABLE I
NEW QUATERNARY CODES

w	length	dimension k	our d	Brouwer's d_B
15	135	45	43	42
13	133	45	42	41
11	131	45	41	40
9	129	45	40	39
7	127	45	39	38
5	125	45	38	37
3	123	45	37	36
2	122	45	36	35
1	121	45	36	35
0	120	45	35	34
12	132	44	42	41
10	130	44	41	40
8	128	44	40	39
6	126	44	39	38
4	124	44	38	37
2	122	44	37	36
0	120	44	36	35

There are totally 44 cosets. Let $\{u_1, \dots, u_{44}\}$ be the maximal complete system of representatives of 5-cyclotomic cosets modulo 124.

Let $q = s = 5$. It is easy to see that for each coset S, we have $\dim_{\boldsymbol{F}_5} V(S_{u_i}) = 1$ by Lemma 3.2.

By taking

$$h(x) = (x^{124} - 1)/(x-1)(x-2)(x-3).$$

Then $h(x) \in \boldsymbol{F}_5[x]$ has one monic linear divisor and 40 distinct irreducible monic divisors of degree 3. Thus, the code $C_{93}(5, h(x))$ is a 5-ary $[41, \geq k, \geq d]$-linear codes with $k = 44-(44-24) = 24$ by Theorem 3.5 (note that $u_{24} = 93$), and

$$d \geq 41 - 1 - \lfloor (93-1)/3 \rfloor = 10$$

by Corollary 2.2. Compared with Brouwer's table [1], the best known minimum distance for a 5-ary $[41, 24]$-linear code is 9. Hence, we get an improvement. From a 5-ary $[41, 24, 10]$-linear code, we can also get a 5-ary $[40, 23, 10]$-linear code. This is also an improvement on Brouwer's table [1].

C. $h(x) = x^N - x$

In this case, we assume that $\gcd(q, N-1) = 1$. Then it is easy to see that $V_{N-1}(x^N - x)$ is generated by $V_{N-2}(x^{N-1} - 1)$ and x^{N-1}. Hence, we have the following result from Theorem 3.5.

Theorem 3.9: Let $C_m(q; x^N - x)$ be the $\boldsymbol{F}_q$-linear code defined in (2.4) with $h(x) = x^N - x$. Let $\{u_1, \dots, u_\ell\}$ be the maximal complete system of representatives of the q-cyclotomic cosets modulo $N-1$ with $0 = u_1 < u_2 < \cdots < u_\ell = N-2$. If r satisfies $u_r \leq m < u_{r+1}$ for some r, then the code has dimension

$$\sum_{i=1}^{r} \dim_{\boldsymbol{F}_q} V(S_{u_i}).$$

In particular, this dimension is at least $n_2 - (\ell - r)\log_q s$, where n_2 is the number of distinct monic irreducible divisors of $x^{N-1} - 1$.

Example 3.10: As in Example 3.8, part ii), we can produce many best known quaternary codes by considering $h(x) = x^{256} - x$. For instance, quaternary $[136, 45, 43]$, $[136, 115, 8]$-linear codes, etc., can be obtained. We skip the detailed computation.

ACKNOWLEDGMENT

We are grateful to Shoulun Long for his help in computing cyclotomic cosets. We also wish to thank San Ling providing us with Remark 3.3 and the referees for their useful comments.

REFERENCES

[1] A. E. Brouwer. Linear Code Bounds. [Online]. Available: http://www.win.tue.nl/~aeb/voorlincod.html

[2] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1998.

[3] S. Ling, H. Niederreiter, and C. P. Xing, "Symmetric polynomials and some good codes," *Finite Fields Their Applic.*, vol. 7, pp. 142–148, 2001.

[4] C. Xing and S. Ling, "A class of linear codes with good parameters," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2184–2188, Sept. 2000.

Coding and Signal Space Diversity for a Class of Fading and Impulsive Noise Channels

Jürgen Häring and A. J. Han Vinck, *Senior Member, IEEE*

***Abstract*—The transmission over the Gaussian mixture noise channel with perfect channel state information at the receiver side is considered. Lower and upper bounds on the achievable pairwise error probability (PEP) are derived for finite and infinite codeword lengths. It is shown that diversity codes, i.e., unitary transforms, can be applied to achieve a diversity gain. A large class of diversity codes is determined for which—if the codeword length is increased—the PEP between any two codewords approaches either zero or the lower bound on the PEP.**

***Index Terms*—Compound channel, diversity, fading channel, impulsive noise, orthogonal frequency-division multiplexing (OFDM), signal space diversity, unitary transform.**

I. INTRODUCTION

In many cases of practical interest, Gaussian mixture noise (GMN) channel models are well suited to statistically describe the impact of noise on a digital communication system. On the GMN channel, the data is corrupted by additive white Gaussian noise (AWGN) with randomly varying noise variance, see [7]. In this correspondence, we assume that the channel is memoryless and that the receiver is provided with perfect channel state information (CSI), i.e., the receiver knows the noise variance. Typically, reliable CSI can be obtained on channels where strong statistical dependencies between consecutive noise samples exist. If the channel state estimator in the receiver is succeeded by

Manuscript received November 26, 2002; revised December 14, 2003. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Lausanne, Switzerland, June/July 2002.

A. J. H. Häring is with Bosch, 70182 Stuttgart, Germany (e-mail: j.haering@web.de).

A. J. H Vinck is with the Institute for Experimental Mathematics, University of Essen, 45326 Essen, Germany (e-mail: vinck@exp-math.uni-essen.de).

Communicated by R. Koetter, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2004.826668

an interleaver of sufficient length, the channel appears memoryless to any further signal processing or decoding algorithm.

In this correspondence, only minor assumptions are made for the statistics of the channel state. Hence, our analysis covers a whole class of communication channels, including several well-known cases; if, e.g., the channel state is constant, the classical AWGN model is obtained; other examples are the slowly, flat Rayleigh-fading channel with perfect CSI, see [11], and the Gaussian collision channel, see [4]; finally, many models for impulsive noise channels are included in the GMN channel model, see [6], [7]. In a simple model, a good channel state with a low noise variance (background noise) and a bad channel state with a high noise variance (impulse) might be defined. Reliable CSI can then be obtained if the noise is, e.g., bursty or the variances in the good and the bad channel state vary significantly.

The correspondence is organized as follows. Section II introduces the GMN channel model. Based on the union bound, the decoding error probability P_e achieved by block-coded transmission and maximum *a posteriori* (MAP) decoding is considered in Section III. The pairwise error probability (PEP) is calculated, and achievable lower bounds on the PEP are derived for both finite and infinite codeword lengths. This discussion is related to the analysis of the Rayleigh-fading channel with perfect CSI in [2], [3], [14]. We, therefore, compare our results to the Rayleigh-fading case and give rules for the design of "good" codes. In Section IV, the concept of so-called *diversity codes* is introduced. The diversity encoder applies a unitary transform over the real or complex numbers to rotate a vector of information symbols in the n-dimensional space, where n is the length of the information sequence. The transform leaves the Euclidean distance between any two information vectors unchanged, however, it modifies the distribution of the Euclidean distance over the components of the information vector and therefore introduces signal space diversity. Section IV discusses a class of diversity codes for which, as n is increased, the PEP of any arbitrarily chosen pair of codewords approaches either zero or the best possible diversity distribution. This supports the observation made in [8] that several diversity codes approach the same performance as n is increased.

Diversity codes were already applied in 1963 by Lang for the transmission over an impulsive noise channel, see [9]. In [2], [8], the construction of good diversity codes is studied for the slowly, flat Rayleigh-fading channel with perfect CSI, whereas [12] and [15] focus on the decoding problem. In [12], a suboptimal decoder based on the decision feedback principle is proposed that is mainly suited for long codeword lengths, see [8]. In contrast, the lattice decoder presented in [15] requires small codeword lengths to operate with acceptable complexity. Finally, it should be noted that also a convolutional type of diversity codes exist, see [16], and that [6] analyzes the performance and decoding of diversity codes for the GMN channel without CSI.

II. Channel Model

At each discrete time instant, the GMN channel accepts one symbol $x \in \mathcal{X}$ from the channel input alphabet $\mathcal{X}$ and maps it onto a symbol $y \in \mathcal{Y}$ of the channel output alphabet $\mathcal{Y}$. Conditioned on the random channel state $s \in \mathcal{S}$, the channel adds white Gaussian noise w with variance σ_s^2

$$y = x + w \tag{1}$$

where $\mathcal{S}$ is the set of all possible channel states, and σ_s^2 is a deterministic function of $s \in \mathcal{S}$. Since we assume perfect CSI, the probability density function (pdf) of w is conditioned on the channel state

$$p(w|s) = \frac{1}{2\pi\sigma_s^2}\exp\left(-\frac{|w|^2}{2\sigma_s^2}\right) \tag{2}$$

where we assume that the channel output alphabet is complex (passband transmission), see, e.g., [7]. Note that the results presented in the following are similar for the channel with real input and output alphabets. Applying (1), the channel transition pdf immediately follows as

$$p(y|x,s) = p(w = y - x|s) = \frac{1}{2\pi\sigma_s^2}\exp\left(-\frac{(|y-x|^2}{2\sigma_s^2}\right). \tag{3}$$

Since the channel is memoryless, the channel state is an independent random variable not depending on any previous or succeeding states. It is completely described by the pdf $p(s)$ and the probability mass function (pmf) $P(s)$, respectively.

For the GMN channels discussed in this correspondence, the following assumptions are made:

- the special case where the GMN channel degenerates to the AWGN channel, i.e., one channel state is taken with probability one, is not considered;
- $\mathrm{E}_s\left\{1/\sigma_s^2\right\} < \infty$, where $\mathrm{E}_s\{\cdot\}$ denotes the expectation over the channel state s;
- only channels providing a finite average noise variance $\sigma_w^2 = \mathrm{E}_s\left\{\sigma_s^2\right\} < \infty$ are considered.

To illustrate the results obtained in this correspondence, we introduce a simple, two-state GMN channel model. The states $s \in \{0,1\}$ are taken with probabilities $P(s=0) = 0.9$, $P(s=1) = 0.1$. The corresponding noise variances $\sigma_{s=0}^2$, $\sigma_{s=1}^2$ directly follow from $\sigma_w^2 = \mathrm{E}_s\left\{\sigma_s^2\right\}$ and the parameter $T := \sigma_{s=0}^2/\sigma_{s=1}^2$, with $T \neq 0$ (for $T = 0$ the model is equivalent to the AWGN channel). For $T \ll 1$, this model describes an impulsive noise channel where 10% of the transmitted symbols are hit by an impulse, i.e., Gaussian noise with a large variance, and 90% are corrupted by background noise with $\sigma_{s=0}^2$. In the sequel, this channel will be referred to as impulsive noise channel.

As another example, the Rayleigh-fading channel with perfect CSI, see, e.g., [1] for a more detailed treatment, is covered by the above GMN model. On this channel, each transmitted symbol x is multiplied with the Rayleigh distributed fading coefficient s, and AWGN with variance σ_w^2 is added. Considering s as the channel state and normalizing the channel output to s, the received samples can be described by the GMN model with $p(s) = 2s\exp(-s^2)$ and $\sigma_s^2 = \sigma_w^2/s^2$.

Finally, we define the signal-to-noise ratio (SNR) as $\mathrm{SNR} := \mathrm{E}_x\left\{|x|^2\right\}/N_0$, where $\mathrm{E}_x\left\{|x|^2\right\}$ is the average power of the transmitted symbols, and $N_0 := 2\sigma_w^2$ is the single-sided noise power spectral density.

III. Decoding Error Probability

Since exactly calculating the decoding error probability P_e is a difficult task, we consider the union upper bound on P_e

$$P_e := P(\boldsymbol{c}' \neq \boldsymbol{c}) \leq \sum_{\boldsymbol{c}\in\mathcal{C}} P(\boldsymbol{c}) \sum_{\substack{\boldsymbol{c}'\in\mathcal{C}\\ \boldsymbol{c}'\neq\boldsymbol{c}}} P(\boldsymbol{c}\to\boldsymbol{c}') \tag{4}$$

where $P(\boldsymbol{c})$ is the probability that codeword $\boldsymbol{c}$ is transmitted, and $P(\boldsymbol{c}\to\boldsymbol{c}')$ is referred to as the PEP. The PEP is the probability that, in the binary decision between $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{c}' \in \mathcal{C}$, $\boldsymbol{c}'$ is erroneously decoded given that $\boldsymbol{c}$ was transmitted. Within this correspondence, we assume that every codeword $\boldsymbol{c} \in \mathcal{C}$ is transmitted equally likely with probability $P(\boldsymbol{c}) = 1/|\mathcal{C}|$. Clearly, the number of codeword pairs achieving the maximum PEP must be small to minimize the bound.

For the further discussion, it is useful to define the distance vector.

Definition III.1 (Distance Vector): For two codewords $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$ of length n, the components $d_k := |c_k - c'_k|$ of the *distance vector*

$\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')=(d_1,\dots,d_n)$ are defined as the Euclidean distance between the code symbols c_k and c'_k.

A. MAP Decoding

The decoding of the transmitted codeword is based on the received vector $\boldsymbol{y}=(y_1,\dots,y_n)$ and the channel state vector $\boldsymbol{s}=(s_1,\dots,s_n)$. The MAP decoder chooses the codeword $\boldsymbol{c}=(c_1,\dots,c_n)$ that maximizes the *a posteriori* probability $P(\boldsymbol{c}|\boldsymbol{y},\boldsymbol{s})$, see [17]. For equally likely transmitted codewords, the MAP decoding rule is given by

$$\underset{\boldsymbol{c}\in\mathcal{C}}{\operatorname{argmax}}\, P(\boldsymbol{c}|\boldsymbol{y},\boldsymbol{s})=\underset{\boldsymbol{c}\in\mathcal{C}}{\operatorname{argmax}}\prod_{k=1}^{n}p(w_k=y_k-c_k|s_k) \tag{5}$$

with the pdf $p(w_k|s_k)$ defined by (2). Further simplifying (5) shows that MAP decoding is equivalent to choosing the codeword that *minimizes* the additive decoding metric

$$\omega(\boldsymbol{c},\boldsymbol{y},\boldsymbol{s}):=\sum_{k=1}^{n}\frac{|y_k-c_k|^2}{\sigma^2_{s_k}}. \tag{6}$$

1) PEP: To evaluate the union upper bound on P_e, see (4), the PEP $P(\boldsymbol{c}\to\boldsymbol{c}')$ has to be evaluated. With the MAP decoding metric $\omega(\boldsymbol{c},\boldsymbol{y},\boldsymbol{s})$ in (6), the PEP conditioned on the channel state vector $\boldsymbol{s}$ is given by

$$P(\boldsymbol{c}\to\boldsymbol{c}'|\boldsymbol{s})=P(\omega(\boldsymbol{c}',\boldsymbol{y},\boldsymbol{s})\le\omega(\boldsymbol{c},\boldsymbol{y},\boldsymbol{s})|\boldsymbol{s}).$$

This equation can be rewritten in the form

$$P(\boldsymbol{c}\to\boldsymbol{c}'|\boldsymbol{s})=P(M\le\Omega|\boldsymbol{s})$$

with the constant $M=\sum_{k=1}^{n}d_k^2/(2\sigma^2_{s_k})$ and the random variable

$$\Omega=\sum_{k=1}^{n}\frac{\operatorname{Re}\{c'_k-c_k\}}{\sigma^2_{s_k}}\operatorname{Re}\{w_k\}+\frac{\operatorname{Im}\{c'_k-c_k\}}{\sigma^2_{s_k}}\operatorname{Im}\{w_k\}$$

where $\operatorname{Re}\{\cdot\}$ and $\operatorname{Im}\{\cdot\}$ denote the real and imaginary parts of their complex argument, respectively. Since for a given channel state s_k the noise w_k is complex Gaussian distributed, see (2), Ω is a real Gaussian distributed random variable with expectation $\mathrm{E}\{\Omega\}=0$ and variance $\mathrm{E}\{\Omega^2\}=2M$. Hence, it follows

$$\begin{aligned}P(\boldsymbol{c}\to\boldsymbol{c}'|\boldsymbol{s})&=\int_{-\infty}^{M}p(\Omega=x|\boldsymbol{s})dx=\frac{1}{2}\operatorname{erfc}\left(\frac{M}{\sqrt{2\mathrm{E}\{\Omega^2\}}}\right)\\&=\frac{1}{2}\operatorname{erfc}\left(\sqrt{\sum_{k=1}^{n}\frac{d_k^2}{8\sigma^2_{s_k}}}\right).\end{aligned} \tag{7}$$

To obtain $P(\boldsymbol{c}\to\boldsymbol{c}')$, the expectation over the channel state vector $\boldsymbol{s}$ has to be evaluated. The following theorem resumes the final result.

Theorem III.1 (PEP Given Perfect CSI): For the GMN channel with perfect CSI adopting MAP decoding, the PEP for the codewords $\boldsymbol{c}$, $\boldsymbol{c}'\in\mathcal{C}$ of length n depends on the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ only and is given by

$$P(\boldsymbol{c}\to\boldsymbol{c}')=\mathrm{E}_{\boldsymbol{s}=(s_1,\dots,s_n)}\left\{\frac{1}{2}\operatorname{erfc}\left(\sqrt{\sum_{k=1}^{n}\frac{d_k^2}{8\sigma^2_{s_k}}}\right)\right\}. \tag{8}$$

The result holds for both, the real and the complex GMN channel with perfect CSI.

Since evaluating the n-dimensional expectation operator given above can be quite complicated, we will frequently use the upper Chernoff bound on the PEP in the following. It is based on the well-known relation $\operatorname{erfc}(x)\le\exp(-x^2)$, see [17], yielding

$$P(\boldsymbol{c}\to\boldsymbol{c}')\le\frac{1}{2}\prod_{k=1}^{n}\mathrm{E}_{s_k}\left\{\exp\left(-\frac{d_k^2}{8\sigma^2_{s_k}}\right)\right\}. \tag{9}$$

Compared to (8), this bound can be quickly evaluated since the expectation is only one-dimensional.

B. PEP Analysis

In this subsection, we study the importance of the diversity between two codewords, i.e., we follow the question how a constant, given Euclidean distance $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ between two codewords should be distributed over the symbols d_k of the distance vector to achieve a low PEP. This will show a major difference between the GMN channel with perfect CSI and the AWGN channel; on the AWGN channel, the PEP can only be decreased by increasing $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$, whereas, on the GMN channel, the diversity can be even more important.

Theorem III.2 considers the case where only two components of the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ are modified while $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ remains unchanged.

Theorem III.2 (Convexity): Assume that all components d_k, $k\in\{1,\dots n\}\setminus\{i,j\}$ of the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ are constant. It is required that $M^2:=d_i^2+d_j^2$ is constant which implies that also $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ is constant. Then, the PEP $P(\boldsymbol{c}\to\boldsymbol{c}')$ is a function of d_i which is minimized for $d_i=M/\sqrt{2}$. Its maximum is achieved for $d_i=0$ and $d_i=M$, respectively. Moreover, applying the substitution $z:=d_i^2$, the PEP is a convex function of z that is symmetrical about its minimum $z=M^2/2$.

The proofs of this and all further theorems can be found in the Appendix.

Example III.1: Fig. 1 depicts the normalized PEP achieved on the impulsive noise and the Rayleigh-fading channel with perfect CSI for $\sigma_w^2=1$ and codeword length $n=2$. The PEP has been normalized to the PEP achieved by

$$\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')=(|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|,0)$$

to be able to show all curves in one plot. The convex structure of the PEP can be clearly observed. For $T=10^{-4}$, the smallest possible PEP is already almost achieved for values of d_1^2 that significantly differ from the optimum $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2/2$. From the standpoint of code design, this property is desirable since increasing the Hamming distance between the codewords is sufficient to almost achieve the whole diversity gain. In contrast, for $T=5\cdot10^{-2}$ and the Rayleigh-fading channel, the figure shows that d_1^2 must be close to 0.5 to (almost) achieve the minimum PEP. Hence, here a large Hamming distance does not guarantee a low PEP.

The convexity property shown earlier is now applied to determine the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ that achieves the minimum PEP for a finite codeword length n.

Theorem III.3 (Lower Bound, Finite n): Consider the PEP $P(\boldsymbol{c}\to\boldsymbol{c}')$ for all possible distance vectors $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ of finite length n under the constraint $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|=M$, where $M>0$ is a constant. Then, $P(\boldsymbol{c}\to\boldsymbol{c}')$ achieves its absolute minimum for $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')=\boldsymbol{d}^{\min}$, where the components of $\boldsymbol{d}^{\min}$ are given by $d_k^{\min}:=M/\sqrt{n}$.

In a similar manner, the following theorem identifies the worst case distance vectors, i.e., the $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ achieving the largest PEP for a given fixed Euclidean distance $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$.

Theorem III.4 (Upper Bound, Finite n): Consider the PEP $P(\boldsymbol{c}\to\boldsymbol{c}')$ for all possible distance vectors $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ of finite length n under the constraint $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|=M$, where M is a constant. Then, $P(\boldsymbol{c}\to\boldsymbol{c}')$ achieves its maximum iff the Euclidean distance is concentrated in one component of $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$, i.e., $d_k=M$ and $d_j=0$ holds for any j, $k\in\{1,\dots n\}$, $j\ne k$.

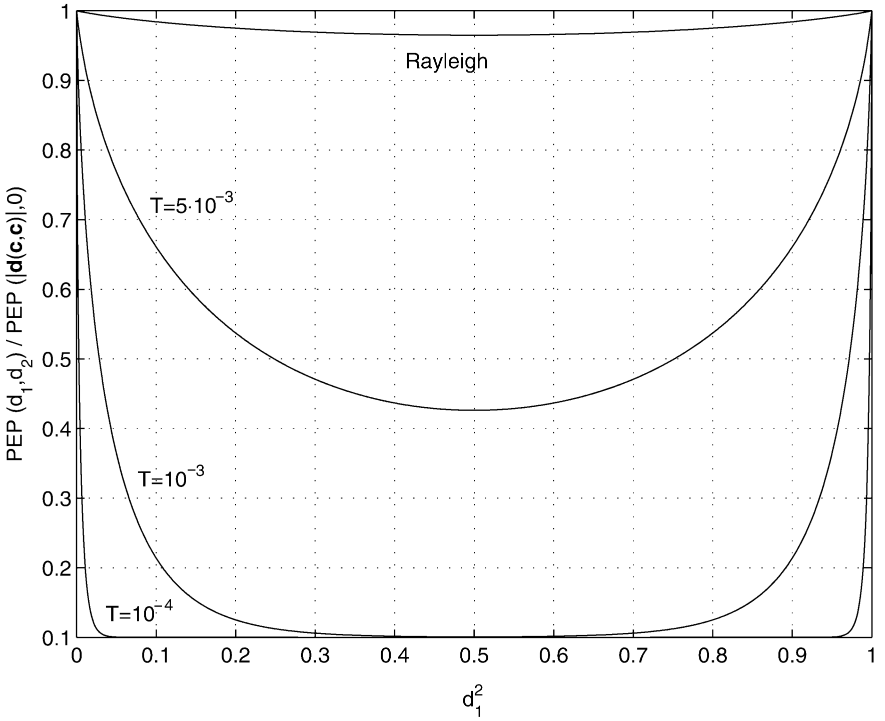

Fig. 1. Normalized PEP achieved by $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}') = (d_1, d_2)$ under the constraint $d_1^2 + d_2^2 = 1$ on the impulsive noise and Rayleigh-fading channel with $\sigma_w^2 = 1$.

The following theorem addresses the question of how far the PEP can be reduced while keeping $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ constant and which vectors $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ achieve this lower bound.

Theorem III.5 (Lower Bound): For a given finite distance $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ between the codewords $\boldsymbol{c},\boldsymbol{c}' \in \mathcal{C}$, the PEP $P(\boldsymbol{c} \rightarrow \boldsymbol{c}')$ is lower-bounded by the PEP achieved on an AWGN channel with variance $1/\mathrm{E}_s\left\{1/\sigma_s^2\right\}$, i.e.,

$$\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2}{8}E_s\left\{\frac{1}{\sigma_s^2}\right\}}\right) \leq P(\boldsymbol{c} \rightarrow \boldsymbol{c}').$$

Equality is achieved in the limiting case $n \rightarrow \infty$ if all $k = 1, \ldots, n$ components d_k of the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ satisfy the condition $\lim_{n\rightarrow\infty}(\sqrt{n}d_k) < \infty$. The theorem holds for all GMN channels providing $\mathrm{E}_s\left\{1/\sigma_s^2\right\} < \infty$ and $\mathrm{E}_s\left\{1/\sigma_s^4\right\} < \infty$.

The preceding theorem shows that a whole class of distance vectors achieves the lower bound for $n \rightarrow \infty$. In contrast, the lower bound for finite n can only be achieved by one specific distance vector, see Theorem III.3.

Finally, it should be noted that all theorems previously given can also be derived for the upper Chernoff bound on the PEP $P(\boldsymbol{c} \rightarrow \boldsymbol{c}')$, see (9).

C. Discussion

The analysis of the PEP as a function of $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ shows that the PEP can be decreased by increasing the Euclidean distance $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ and/or the diversity between the codewords. To achieve a diversity gain, the convexity stated by Theorem III.2 shows that "balanced" distance vectors should be used, and especially zero elements in $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ should be avoided. To make more specific statements—especially if increasing diversity or Euclidean distance leads to higher gains—the particular GMN channel under investigation, the SNR, and also the codeword length have to be further specified. For different GMN channels totally different results might be obtained as demonstrated by the following three examples.

1) Impulsive Noise Channel With $T \approx 1$: For $T = 1$, the impulsive noise channel reduces to the AWGN case. Hence, as $T \rightarrow 1$, diversity becomes less relevant and Euclidean distance dominates the PEP's behavior.

2) Impulsive Noise Channel With $T \ll 1$: For $T \ll 1$, the PEP shows a behavior similar to the symbol error rate (SER) curves in Fig. 3. The error floor typical for impulsive noise channels is introduced by the "bad" channel state $s = 1$ with $\sigma_{s=1}^2 \gg \sigma_{s=0}^2$. This can also be studied analytically: if all $d_k \neq 0$ satisfy $d_k \gg \sigma_{s=0}^2$ (this is true if $0 < M < d_k$ holds for all $d_k \neq 0$ with some finite constant M and sufficiently large SNR), the expectation of all terms with $d_k \neq 0$ in (9) is dominated by $\exp\left(-d_k^2/(8\sigma_{s_k=1}^2)\right)$. Then, for finite n the PEP might be approximated by

$$\begin{aligned} P(\boldsymbol{c} \rightarrow \boldsymbol{c}') &\lessapprox \frac{1}{2}P(s_k = 1)^l \prod_{k=1}^{n} \exp\left(-\frac{d_k^2}{8\sigma_{s_k=1}^2}\right) \\ &= \frac{1}{2}P(s_k = 1)^l \exp\left(-\frac{|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2}{8\sigma_{s_k=1}^2}\right) \end{aligned} \qquad (10)$$

where l is the Hamming weight of $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$, and "$\lessapprox$" can be replaced by "$\leq$" for SNR $\rightarrow \infty$. Note that Theorem III.5 shows that this approximation is not necessarily valid for $l \rightarrow \infty$. It can be observed that increasing the Hamming distance l between two codewords clearly decreases the PEP. Especially in the region of the error floor where $\sigma_{s_k=0}^2 \ll |\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2 \ll \sigma_{s_k=1}^2$ and therefore,

$$\exp\left(-|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2/(8\sigma_{s_k=1}^2)\right) \approx 1$$

holds, increasing l (diversity) is more important than increasing $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ (Euclidean distance). In contrast, for $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2 \gg \sigma_{s_k=1}^2$ (large SNR), $\exp\left(-|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2/(8\sigma_{s_k=1}^2)\right)$ decreases rapidly so that the gain in SNR achieved by diversity converges toward zero and is easily outperformed by a slight increase of $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$.

3) Rayleigh-Fading Channel With Perfect CSI: The Rayleigh-fading channel with perfect CSI has been extensively discussed in literature [1]–[3], [13], [14]. The PEP can be evaluated exactly, see [14], however, rules for code design are usually based on the Chernoff

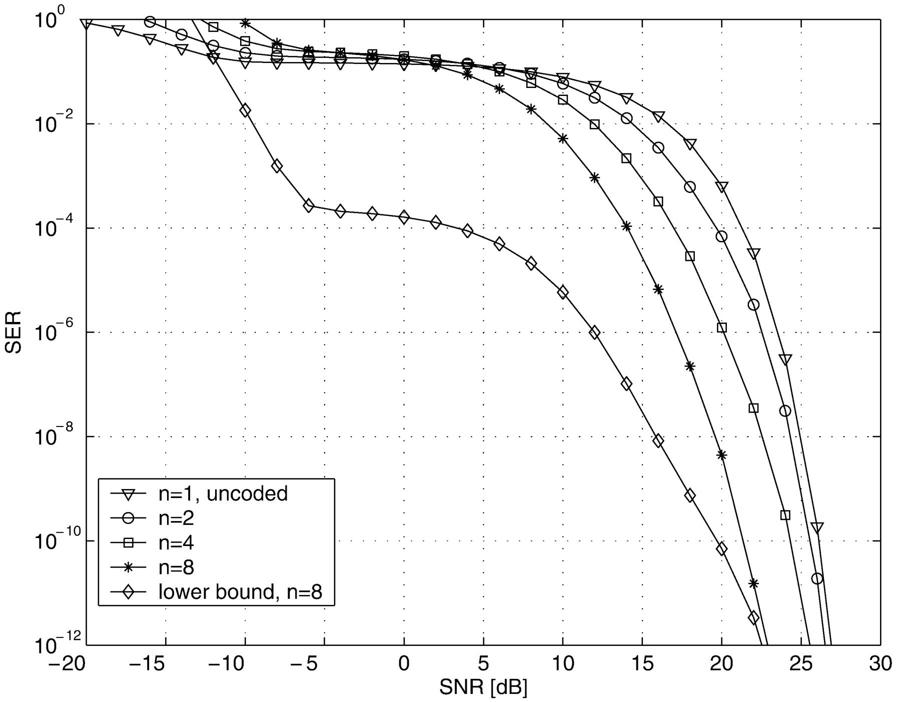

Fig. 2. Union bound for the impulsive noise channel with perfect CSI ($T = 10^{-3}$) using the Chernoff bound on the PEP, $\boldsymbol{G} := \boldsymbol{F}^{-1}$, and 4-QAM modulation.

upper bound on the PEP, see (9). Solving (9) for the Rayleigh-fading channel as defined in Section II yields

$$P(\boldsymbol{c} \to \boldsymbol{c}') \leq \prod_{k=1}^{n} \frac{1}{1 + d_k^2/(8N_0)} \leq \prod_{d_k \neq 0} \frac{1}{d_k^2/(8N_0)} = \frac{1}{(\mathrm{SNR}/8)^l d_p^{(l)}} \tag{11}$$

where the second upper bound is tight for $\min_k d_k^2 N_0 \gg 1$. In the above equation, l again denotes the Hamming weight (or diversity) of $\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')$, and $d_p^{(l)} : \prod_{d_k \neq 0} d_k^2 / \mathrm{E}_x \left\{ |x_k|^2 \right\}$ is defined as the normalized product distance.

The bound shows the following behavior under the assumption $\min_k d_k^2 N_0 \gg 1$ which is typically true in the high SNR domain, see also [3].

- As the SNR increases, diversity l becomes more important; especially for a good asymptotic performance, l is the most important code design parameter.
- At a given diversity l, the product distance $d_p^{(l)}$ should be high. This is achieved by using "balanced" distance vectors or increasing Euclidean distance $|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')|$ between the codewords.

These rules include and further specify the statements made in Section III-C for the design of good codes for GMN channels: the importance of l and the product distance both are consequences of Theorem III.2. However, the results are different as for the impulsive noise channel discussed before. Hence, the well-known code design rules for the Rayleigh-fading channel do not apply to the whole class of GMN channels.

IV. Diversity Block Codes

Input to the diversity encoder is the information vector $\boldsymbol{u} = (u_1, \ldots, u_n) \in \mathcal{U}^n$ of length n. The symbol alphabet $\mathcal{U}$ is a set of $|\mathcal{U}|$ complex numbers defined by the modulation scheme adopted. As an example, for 4-QAM, $\mathcal{U} = \{\pm(1 + j); \pm(1 - j)\}$ holds. Each vector $\boldsymbol{u}$ is encoded by a diversity block code $\mathcal{C}$ with codewords $\boldsymbol{c} = (c_1, \ldots c_n)$ and complex code symbols $c_i \in \mathbb{C}$. The encoding operation is defined by the linear mapping

$$\boldsymbol{c} = \boldsymbol{G}\boldsymbol{u} \tag{12}$$

where $\boldsymbol{G}$ is the unitary $n \times n$ generator matrix. Unitary means that $\boldsymbol{G}^*\boldsymbol{G} = \boldsymbol{I}$ holds, where $\boldsymbol{I}$ is the identity matrix and the asterisk denotes transpose complex conjugate. Possible choices for $\boldsymbol{G}$ are, e.g., the discrete Fourier-, the discrete cosine-, or the Walsh–Hadamard-transform matrix. The codewords $\boldsymbol{c}$ are simply rotated versions of the information vectors $\boldsymbol{u}$ in the n-dimensional space, and, from the Parseval theorem, it follows that the Euclidean distance between two information vectors $\boldsymbol{u}, \boldsymbol{u}'$ is the same as between the codewords $\boldsymbol{c} = \boldsymbol{G}\boldsymbol{u}$, $\boldsymbol{c}' = \boldsymbol{G}\boldsymbol{u}'$

$$|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')| = \sqrt{(\boldsymbol{c} - \boldsymbol{c}')^*(\boldsymbol{c} - \boldsymbol{c}')} = \sqrt{(\boldsymbol{u} - \boldsymbol{u}')^*\boldsymbol{G}^*\boldsymbol{G}(\boldsymbol{u} - \boldsymbol{u}')} = |\boldsymbol{d}(\boldsymbol{u}, \boldsymbol{u}')|. \tag{13}$$

This property explains the name "diversity codes"; whereas "classical" codes designed for the AWGN channel are used to increase the Euclidean distance between the information sequences, the diversity encoder can only increase diversity between the codewords, i.e., distribute a given Euclidean distance over as many code symbols as possible. The aim is to approach the optimum distribution defined by Theorem III.3.

To design codes with both, a high Euclidean distance and good diversity, a product encoder might be applied where the "classical" outer code maximizes the Euclidean distances among the codewords. The inner diversity code is then used to optimize the diversity properties of the outer code while leaving the Euclidean distances unchanged, see also [3], [6].

Example IV.1: To show that the concept of diversity codes is reasonable, Fig. 2 shows the union bound on the SER for a 4-QAM modulation alphabet $\mathcal{U}$ and $\boldsymbol{G} := \boldsymbol{F}^{-1}$, where $\boldsymbol{F}$ is the Fourier matrix. The bounds are plotted for the impulsive noise channel ($T = 10^{-3}$), and, additionally, the minimum possible union bound for $n = 8$ is depicted. It is constructed by assuming that the PEP between all codewords achieves the lower bound on the PEP given by Theorem III.3. Note that we do not state that a code with these properties exists. All curves depicted are based on the upper Chernoff bound, see (9), rather than the exact PEP.

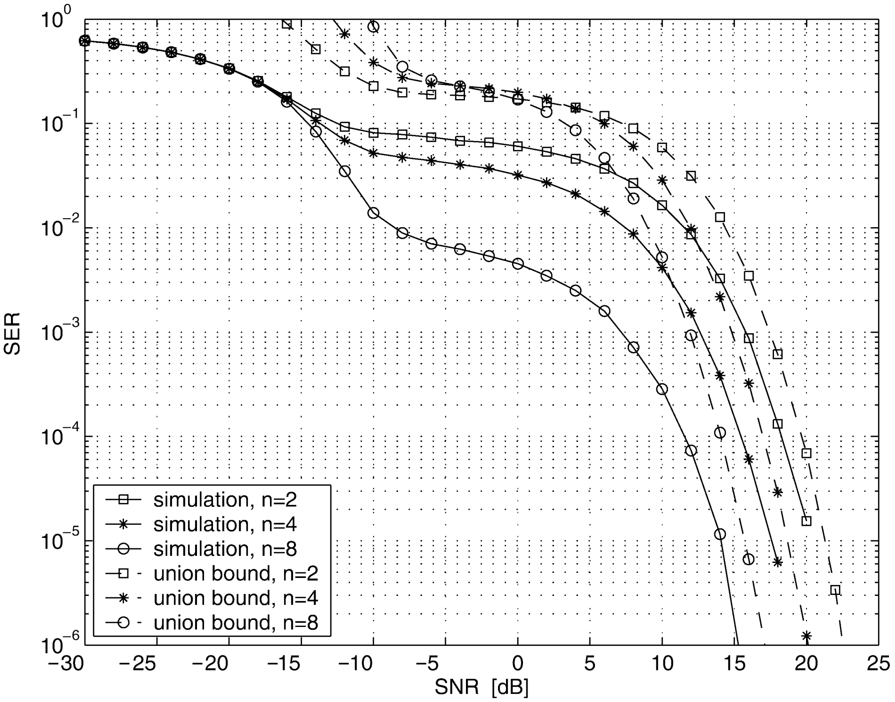

Fig. 3. Union bound and simulation results for the complex impulsive noise channel ($T = 10^{-3}$) using the Chernoff bound on the PEP, $\boldsymbol{G} := \boldsymbol{F}^{-1}$, and 4-QAM modulation.

The figure shows that the union bound decreases with increasing codeword length n and that for SERs of practical interest large diversity gains are achieved. In contrast, in the high-SNR domain, it can be observed that the gaps between the lower union bound, the codes of any length n, and even uncoded transmission become smaller. This is explained by (10) showing that for SNR $\to \infty$, the maximum PEP codewords of every finite-length diversity code have minimum Euclidean distance. These maximum PEP codewords dominate the union bound's behavior for SNR $\to \infty$. Since the minimum Euclidean distance is the same for all codeword lengths n and the lower bound, see (13), the gain in SNR converges toward zero. This holds for every finite-length diversity code with the same modulation alphabet $\mathcal{U}$.

Note that, in contrast, on the Rayleigh-fading channel with perfect CSI diversity codes can achieve a gain in SNR also for SNR $\to \infty$, see Section III-C-III and [2], [6], [8].

Example IV.2: The coding scheme defined in Example IV.1 is now applied to the impulsive noise channel with $T = 10^{-3}$. Fig. 3 compares the union bound on the symbol error rate with simulation results. It can be observed that, for small SERs, the gaps between the union bounds and the simulation results become small. Therefore, in the high-SNR region and for this example, the union bound is a suitable performance measure. We numerically verified for $n = 4$ that the remaining gap results from employing (9) instead of the exact $P(\boldsymbol{c} \to \boldsymbol{c}')$ to compute the union bound.

A. Maximum PEP Codewords

The following theorems study the PEPs of a large class of diversity codes and give insight into the behavior of diversity codes as the codeword length n is increased.

Theorem IV.1 (PEP, Perfect CSI): The set of channels providing $\mathrm{E}_s\left\{1/\sigma_s^2\right\} < \infty$ and $\mathrm{E}_s\left\{1/\sigma_s^4\right\} < \infty$ is considered. The diversity block code $\mathcal{C}$ is defined by the encoding operation $\boldsymbol{c} = \boldsymbol{G}\boldsymbol{u}$. Let any two components $u, u' \in \mathcal{U}$ of the information vector $\boldsymbol{u}$ have a finite Euclidean distance, i.e., $0 < M_1 \leq |u - u'|$ with some constant M_1. For the unitary $n \times n$ generator matrix $\boldsymbol{G} = [g_{k,l}]$, it is assumed that $\lim_{n\to\infty} \sqrt{n}|g_{k,l}| \leq M_0 < \infty$ holds for all $g_{k,l}$ and some constant M_0. Then, for $n \to \infty$, the PEP between any codewords $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$, $\boldsymbol{c} \neq \boldsymbol{c}'$ with finite Euclidean distance $|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')| \leq M_2 < \infty$ achieves the lower bound given by Theorem III.5

$$P(\boldsymbol{c} \to \boldsymbol{c}') = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')|^2}{8}\mathrm{E}_s\left\{\frac{1}{\sigma_s^2}\right\}}\right).$$

This theorem is now applied to identify the largest PEP distance vector in a class of diversity codes for $n \to \infty$:

Theorem IV.2 (Maximum PEP, Perfect CSI): The set of channels providing $\mathrm{E}_s\left\{1/\sigma_s^2\right\} < \infty$ and $\mathrm{E}_s\left\{1/\sigma_s^4\right\} < \infty$ is considered. The diversity block code $\mathcal{C}$ is defined by the encoding operation $\boldsymbol{c} = \boldsymbol{G}\boldsymbol{u}$. Let any two components $u, u' \in \mathcal{U}$ of the information vector $\boldsymbol{u}$ have a finite Euclidean distance, i.e., $0 < M_1 \leq |u - u'|$ with some constant M_1. For the unitary $n \times n$ generator matrix $\boldsymbol{G} = [g_{k,l}]$, it is assumed that $\lim_{n\to\infty} \sqrt{n}|g_{k,l}| \leq M_0 < \infty$ holds for all $g_{k,l}$ and some constant M_0. Then, for $n \to \infty$, the maximum PEP distance vector for any codeword pair $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$, $\boldsymbol{c} \neq \boldsymbol{c}'$ is given by

$$\max_{\boldsymbol{c},\boldsymbol{c}'\in\mathcal{C},\boldsymbol{c}\neq\boldsymbol{c}'} P(\boldsymbol{c} \to \boldsymbol{c}') = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')|^2_{\min}}{8}E_s\left\{\frac{1}{\sigma_s^2}\right\}}\right)$$

where

$$|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')|_{\min} = \min_{\boldsymbol{c},\boldsymbol{c}'\in\mathcal{C},\boldsymbol{c}\neq\boldsymbol{c}'} |\boldsymbol{c} - \boldsymbol{c}'|$$

is the minimum Euclidean distance of $\mathcal{C}$.

As an example, the two previous theorems can be applied when $\boldsymbol{G}$ is the discrete Fourier or the Walsh–Hadamard matrix. The theorems show that all codeword pairs with finite Euclidean distance approach the lower bound on the PEP as the codeword length is increased. This is interesting since two uncoded vectors with a small Euclidean distance also always have a small diversity, leading to the bad performance of uncoded transmission. The diversity encoder removes this disadvantageous property and therefore achieves a diversity gain. The fact that the above theorems hold for every diversity code defined by a—loosely speaking—nonsparse generator matrix $\boldsymbol{G}$ and large n supports the observation made in [8] for several randomly selected and algebraically constructed diversity codes that all show a similar performance for large n.

V. Conclusion

This correspondence analyzes the memoryless GMN channel with perfect CSI. Under only minor assumptions for the statistics of the channel state, lower and upper bounds on the achievable PEP are derived for both finite and infinite codeword lengths. It is shown that, depending on the particular GMN channel under investigation, both Euclidean distance and diversity play an important role in code design. Then, the concept of diversity codes is introduced. For a large class of diversity codes with nonsparse generator matrices it is shown that, when increasing n, all codewords with finite Euclidean distance approach the lowest possible PEP that can be achieved by diversity encoding, i.e., by rotating the codewords in the n-dimensional space while keeping their Euclidean distances constant. This illustrates that the diversity encoder removes the disadvantageous combination of a small Euclidean distance and a small diversity occurring with uncoded transmission. Therefore, applying the codes under investigation to the GMN channel leads to a diversity gain.

Appendix

A. Proof of Theorem III.2

To simplify the notation, only the case is considered where the channel state s_k is a discrete random variable with $s_k \in \{0, 1, \ldots\}$. An extension of the proof to continuous state vectors is straightforward. Without loss of generality, $i = 1$ and $j = 2$ is chosen. The formula for $P(\boldsymbol{c} \rightarrow \boldsymbol{c}')$ given in Theorem III.1 is rewritten by expanding the expectation operator over the channel states s_1, s_2

$$P(\boldsymbol{c} \rightarrow \boldsymbol{c}') = E_{s_3,\ldots s_n}\left\{\sum_{l=0}^{\infty} P(s_k = l)^2 \alpha(d_1, d_2) + \sum_{m=l+1}^{\infty} P(s_k = l)P(s_k = m)\beta(d_1, d_2)\right\} \tag{14}$$

where the abbreviations

$$\alpha(d_1, d_2) := \operatorname{erfc}\left(\sqrt{\varphi + \frac{d_1^2 + d_2^2}{8\sigma_l^2}}\right)$$

$$\beta(d_1, d_2) := \operatorname{erfc}\left(\sqrt{\varphi + \frac{d_1^2}{8\sigma_l^2} + \frac{d_2^2}{8\sigma_m^2}}\right) + \operatorname{erfc}\left(\sqrt{\varphi + \frac{d_0^2}{8\sigma_l^2} + \frac{d_1^2}{8\sigma_{m_0}^2}}\right)$$

and

$$\varphi := \sum_{k=3}^{\infty} d_k^2 / 8\sigma_{s_k}^2$$

are introduced. In the following, all parameters other than d_1, d_2 are treated as constants. Applying the condition $M^2 = d_1^2 + d_2^2 = \text{const}$ shows that $\alpha(d_1, d_2)$ is only a function of M^2 and therefore constant. Substituting with $z := d_1^2$ and $M^2 = d_1^2 + d_2^2$ in $\beta(d_1, d_2)$ yields the function $\tilde{\beta}(z, M)$. Straightforward computation shows that

$$\tilde{\beta}(M^2/2 - x, M) = \tilde{\beta}(M^2/2 + x, M)$$

holds for $0 \leq x \leq M^2/2$, i.e., $\tilde{\beta}(z, M)$ is symmetrical to $M^2/2$. Moreover, since the condition $\frac{\partial^2}{\partial y^2}\tilde{\beta}(z, M) > 0$ holds for $\sigma_l^2 \neq \sigma_m^2$ and $\tilde{\beta}(z, M)$ is continuous, it follows that $\tilde{\beta}(z, M)$ is a convex function in z. Both the convexity and the symmetry show that $\tilde{\beta}(z, M)$ achieves its absolute minimum for $z = M^2/2$ and its maxima for $z = 0$ and $z = M^2$, respectively. This implies that $\beta(d_1, d_2)$ achieves its absolute minimum for $d_1^2 = M^2/2$ and its maxima at $d_1 = 0$ and $d_1 = M$, respectively, and that no other extremum exists.

The PEP given by (14) is calculated by summing over the $\alpha(d_1, d_2)$ and $\beta(d_1, d_2)$. Since $\alpha(d_1, d_2)$ is a constant and the $\beta(d_1, d_2)$ achieve their maxima and minima all for the same d_1, d_2, also the PEP achieves its maxima and minima in these points. This proves the theorem. □

B. Proof of Theorem III.3

The set Ω of all distance vectors with Euclidean distance $|\boldsymbol{d}(\boldsymbol{c}, \boldsymbol{c}')|$ is defined as

$$\Omega := \left\{(d_1, \ldots d_n) \,\middle|\, M^2 = \sum_{k=1}^{n} d_k^2 \right\}.$$

Consider a vector $\boldsymbol{d}^{(l=1)} \in \Omega$, $\boldsymbol{d}^{(l=1)} \neq \boldsymbol{d}^{\min}$ different from the minimum as proposed by the theorem. Theorem III.2 ensures that a vector $\boldsymbol{d}^{(l+1)} \in \Omega$ can always be derived from $\boldsymbol{d}^{(l)}$ which achieves a PEP smaller than $\boldsymbol{d}^{(l)}$. This is done by applying the following algorithm: choose an arbitrary value $i = 1, \ldots, n-1$ for which the components of $\boldsymbol{d}^{(l)}$ satisfy $d_i^{(l)} \neq d_{i+1}^{(l)}$. Then, the components of $\boldsymbol{d}^{(l+1)}$ are chosen as

$$d_i^{(l+1)} = d_{i+1}^{(l+1)} := \sqrt{\left({d_i^{(l)}}^2 + {d_{i+1}^{(l)}}^2\right)/2}$$

and $d_k^{(l+1)} = d_k^{(l)}$ for all $k \neq i, i+1$. If $\boldsymbol{d}^{(l+1)} \neq \boldsymbol{d}^{\min}$ holds, the algorithm is applied again to $\boldsymbol{d}^{(l+1)}$. Since in each iteration the PEP is deceased, we refer to the above algorithm as *downhill algorithm*. It is obvious that after infinitely many iterations $\lim_{l\rightarrow\infty} \boldsymbol{d}^{(l)} \boldsymbol{d}^{\min}$ is achieved. Since the algorithm always converges into $\boldsymbol{d}^{\min}$ independently of the starting vector, $\boldsymbol{d}^{\min}$ is the absolute minimum.

For the formal proof that the downhill algorithm converges into $\boldsymbol{d}^{\min}$, the algorithm is described by ${\boldsymbol{d}^{(l+1)}}^2 = \boldsymbol{\Lambda_i} {\boldsymbol{d}^{(l)}}^2$, with ${\boldsymbol{d}^{(l)}}^2 := ({d_1^{(l)}}^2, \ldots {d_n^{(l)}}^2)$ and the $n \times n$ matrix $\boldsymbol{\Lambda_i}$ with elements $\lambda_{i,i} = \lambda_{i+1,i} = \lambda_{i,i+1} = \lambda_{i+1,i+1} = 0.5$, $\lambda_{k,k} = 1$, $k \neq i$, $i+1$, and all other $\lambda_{k,j} = 0$. When applying the algorithm consecutively for $i = 1, 2, \ldots, n-1$ and repeating this for m times, the result is give by $\boldsymbol{\Lambda}^m {\boldsymbol{d}^{(l)}}^2$ with $\boldsymbol{\Lambda} := \prod_{i=1}^{n-1} \boldsymbol{\Lambda_i}$. Evaluating the eigenvalues and vectors of $\boldsymbol{\Lambda}$ shows that for $m \rightarrow \infty$ each element in $\boldsymbol{\Lambda}^m$ equals $1/n$, and therefore $\lim_{l\rightarrow\infty} \boldsymbol{d}^{(l)} = \boldsymbol{d}^{\min}$ holds. □

C. Proof of Theorem III.4

By converting the "downhill algorithm" applied in the proof of Theorem III.3 into an "uphill algorithm," the theorem immediately follows. □

D. Proof of Theorem III.5

The following derivation of the PEP is similar to [2]. The PEP $P(\boldsymbol{c} \rightarrow \boldsymbol{c}')$, see (8), is rewritten in the form

$$P(\boldsymbol{c} \rightarrow \boldsymbol{c}') = \lim_{n\rightarrow\infty} \mathrm{E}_s\{P(\boldsymbol{c} \rightarrow \boldsymbol{c}' | \boldsymbol{s})\} = \frac{1}{2} \lim_{n\rightarrow\infty} \int_{-\infty}^{\infty} \operatorname{erfc}\left(\sqrt{\frac{1}{8}x}\right) p(X_n = x)dx \tag{15}$$

where the definition $X_n := \sum_{k=1}^{n} d_k^2/\sigma_{s_k}^2$ is introduced. In the following, a lemma according to Helly, see [5], will be employed. The lemma states that

$$\lim_{n\rightarrow\infty} \int_{-\infty}^{\infty} g(x)dF_n(x) = \int_{-\infty}^{\infty} g(x)dF_\infty(x)$$

where $F_n(x)$ and $F_\infty(x)$ are the probability distributions of X_n and $\lim_{n\rightarrow\infty} X_n$, respectively, and $g(x)$ is some bounded, continuous function. Note that the Stieltjes integral is used in the above equation.

However, in the following, the less formal description using pdfs and the delta distribution to jointly describe discrete and continuous random variables is employed. Applying Helly's lemma to (15) yields

$$P(\boldsymbol{c} \to \boldsymbol{c}') = \frac{1}{2}\int \operatorname{erfc}\left(\sqrt{\frac{1}{8}x}\right) p(X_\infty = x)dx \tag{16}$$

where the notation $p(X_\infty) := \lim_{n\to\infty} p(X_n)$ is used. $p_{X_\infty}(\cdot)$ is calculated by rewriting X_n in the form

$$X_n = \frac{1}{n}\sum_{k=1}^{n} (\sqrt{n}d_k)^2/\sigma_{s_k}^2 =: \frac{1}{n}\sum_{k=1}^{n} Y_k.$$

The strong law of large numbers states

$$P\left(\lim_{n\to\infty}\left|X_n - \frac{1}{n}\sum_{k=1}^{n}\mathrm{E}_{s_k}\{Y_k\}\right| = 0\right) = 1$$

and according to Kolmogoroff it can be applied if the condition $\sum_{k=0}^{\infty}\sigma_{Y_k}^2/k^2 < \infty$ holds, see, e.g., [5], where the variance $\sigma_{Y_k}^2$ of the random variable Y_k is given by

$$\sigma_{Y_k}^2 = n^2 d_k^4\left(\mathrm{E}_{s_k}\left\{1/\sigma_{s_k}^4\right\} - \mathrm{E}_{s_k}\left\{1/\sigma_{s_k}^2\right\}\right).$$

From the requirements $\lim_{n\to\infty}(\sqrt{n}d_k) \le M_0 < \infty$, $\mathrm{E}_s\left\{1/\sigma_s^4\right\} \le M_1 < \infty$, and $\mathrm{E}_s\left\{1/\sigma_s^2\right\} \le M_2 < \infty$ with the constants M_0, M_1, M_2 it follows for Kolmogoroff's condition

$$\sum_{k=0}^{\infty}\frac{1}{k^2}(\mathrm{E}_{s_k}\{Y_k^2\} - \mathrm{E}_{s_k}\{Y_k\}^2) \le M_0^4(M_1 - M_2)\sum_{k=0}^{\infty}\frac{1}{k^2} < \infty.$$

This means that the condition is fulfilled and therefore the strong law of large numbers can be applied, i.e., $p(X_\infty) = \delta(X_\infty - M)$ with

$$M := \frac{1}{n}\sum_{k=1}^{n}\mathrm{E}_{s_k}\{Y_k\}|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2\mathrm{E}_{s_k}\left\{\frac{1}{\sigma_{s_k}^2}\right\}$$

and the Dirac distribution $\delta(\cdot)$. Utilizing $p(X_\infty)$, (16) can be solved. This yields the bound $P(\boldsymbol{c} \to \boldsymbol{c}') = 0.5\cdot\operatorname{erfc}(\sqrt{M/8})$ as stated in the theorem.

Finally, it has to be shown that the above bound is a lower bound. From Theorem III.3, it is known that $P(\boldsymbol{c} \to \boldsymbol{c}')$ is minimized for $n \to \infty$ if $d_k := |\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|/\sqrt{n}$ holds. This vector also fulfills the requirements of Theorem III.5 and achieves the bound as stated above. Hence, the bound given in the statement of the theorem is indeed a lower bound. □

E. Proof of Theorem IV.1

To be able to apply Theorem III.5, for each component $d_k = |c_k - c_k'|$ of the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ the condition

$$\lim_{n\to\infty}\sqrt{n}d_k < \infty \tag{17}$$

must be fulfilled. In the following, it is shown that this condition holds which completes the proof. We define the vector $\boldsymbol{\Delta} := \boldsymbol{u} - \boldsymbol{u}'$, where $\boldsymbol{u}, \boldsymbol{u}'$ are two information vectors with components $u_k, u_k' \in \mathcal{U}$. The mappings $\boldsymbol{c} = \boldsymbol{G}\boldsymbol{u}$, $\boldsymbol{c}' = \boldsymbol{G}\boldsymbol{u}'$ assign a codeword to each of the information vectors. Then, applying the condition given by (17) to the components of the distance vector $\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')$ yields

$$\begin{aligned}\lim_{n\to\infty}\sqrt{n}d_k &= \lim_{n\to\infty}\sqrt{n}\left|\sum_{l=1}^{n} g_{k,l}\Delta_l\right| \\ &\le \lim_{n\to\infty}\sum_{l=1}^{n}\sqrt{n}|g_{k,l}||\Delta_l| \\ &= \lim_{n\to\infty}\sum_{\{l|\Delta_l\neq 0\}}\sqrt{n}|g_{k,l}||\Delta_l|. \end{aligned}\tag{18}$$

We denote the number of components $\Delta_l \neq 0$ by $|\{l|\Delta_l \neq 0\}|$. It is upper-bounded by $|\{l|\Delta_l \neq 0\}| \le |\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2/M_1^2 \le \infty$. Moreover, since the generator matrix is unitary, $|\Delta_l|^2 \le |\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2$ holds. Substituting with these bounds and the assumption $\lim_{n\to\infty}\sqrt{n}|g_{k,l}| \le M_0 < \infty$ in (18) yields

$$\lim_{n\to\infty}\sqrt{n}d_k \le \frac{|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2}{M_1^2}M_0|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|^2 < \infty$$

where this term is finite since M_0 and M_1 are constants independent of n. □

F. Proof of Theorem IV.2

For $n \to \infty$, the set of all possible distance vectors in the code is divided into the set of vectors with finite and infinite $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$, respectively. Theorem IV.1 shows that, for $n \to \infty$, the PEP of all distance vectors with finite $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ only depends on $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$. Clearly, within this set, the maximum PEP is achieved by the distance vector with the minimum Euclidean distance $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|_{\min}$

$$P(\boldsymbol{c} \to \boldsymbol{c}') \le \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|_{\min}^2}{8}E_s\left\{\frac{1}{\sigma_s^2}\right\}}\right). \tag{19}$$

If $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ is not finite, employing the upper bound on the PEP given by Theorem III.4 yields

$$P(\boldsymbol{c} \to \boldsymbol{c}') \le \mathrm{E}_s\left\{\frac{1}{2}\operatorname{erfc}\left(\frac{|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|}{\sqrt{8\sigma_s^2}}\right)\right\}.$$

For any arbitrary fixed noise variance $\sigma_w^2 = \mathrm{E}_s\left\{\sigma_s^2\right\}$, this bound is a strictly monotonic decreasing function of $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')|$ approaching zero for $|\boldsymbol{d}(\boldsymbol{c},\boldsymbol{c}')| \to \infty$. Hence, this bound, and therefore also the real PEP is always smaller than the PEP given by (19). It follows that, for $n \to \infty$, the maximum PEP in $\mathcal{C}$ is given by (19) which proves the statement of the theorem. □

REFERENCES

[1] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[2] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453–1467, July 1998.

[3] J. Boutros, E. Viterbo, C. Rastello, and J. C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 502–518, Mar. 1996.

[4] C. Caire, E. Leonardi, and E. Viterbo, "Modulation and coding for the Gaussian collision channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2007–2026, Sept. 2000.

[5] M. Fisz, *Wahrscheinlichkeitsrechnung und Mathematische Statistik* (in German), 7th ed. Berlin, Germany: VEB Deutscher Verlag der Wissenschaft, 1973.

[6] J. Häring, *Error Tolerant Communication Over the Compound Channel*. Aachen, Germany: Shaker-Verlag, 2002.

[7] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. Berlin, Germany: Springer-Verlag, 1988.

[8] C. Lamy and J. Boutros, "On random rotations diversity and minimum MSE decoding of lattices," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1584–1589, July 2000.

[9] G. R. Lang, "Rotational transformation of signals," *IEEE Trans. Inform. Theory*, vol. IT-9, pp. 191–197, July 1963.

[10] D. Middleton, "Canonical and quasicanonical probability models of class A interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-25, pp. 76–106, May 1983.

[11] J. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[12] M. Reinhard and J. Lindner, "Transformation of a Rayleigh fading channel into a set of parallel AWGN channels and its advantage for coded transmission," *Electron. Lett.*, vol. 31, no. 25, pp. 2154–2155, Dec. 1995.

[13] C. Schlegel and D. J. Costello, "Bandwidth efficient coding for fading channels: Code construction and performance analysis," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1356–1368, Dec. 1989.

[14] G. Taricco and E. Viterbo, "Performance of high-diversity multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1539–1543, July 1998.

[15] E. Viterbo and J. Boutros, "An universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1639–1642, July 1999.

[16] G. W. Wornell, "Spread-response precoding for communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 488–501, Mar. 1996.

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

On Multicarrier Signals Where the PMEPR of a Random Codeword is Asymptotically $\log n$

Masoud Sharif, *Student Member, IEEE,* and Babak Hassibi

***Abstract*—Multicarrier signals exhibit a large peak-to-mean envelope power ratio (PMEPR). In this correspondence, without using a Gaussian assumption, we derive lower and upper probability bounds for the PMEPR distribution when the number of subcarriers n is large. Even though the worst case PMEPR is of the order of n, the main result is that the PMEPR of a random codeword $C = (c_1, \ldots, c_n)$ is $\log n$ with probability approaching one asymptotically, for the following three general cases: i) c_i's are independent and identically distributed (i.i.d.) chosen from a complex quadrature amplitude modulation (QAM) constellation in which the real and imaginary part of c_i each has i.i.d. and even distribution (not necessarily uniform), ii) c_i's are i.i.d. chosen from a phase-shift keying (PSK) constellation where the distribution over the constellation points is invariant under $\pi/2$ rotation, and iii) C is chosen uniformly from a complex sphere of dimension n. Based on this result, it is proved that asymptotically, the Varshamov–Gilbert (VG) bound remains the same for codes with PMEPR of less than $\log n$ chosen from QAM/PSK constellations.**

***Index Terms*—Multicarrier signals, orthogonal frequency-division multiplexing (OFDM), peak-to-mean envelope power ratio (PMEPR), spherical codes, symmetric constellations.**

I. Introduction

Multicarrier modulation has been proposed in different broad-band wireless and wireline applications such as wireless local area networks (WLAN) and digital subscriber line (DSL). Even though multicarrier modulation has a nice performance in a multipath fading environment, it suffers from high amplitude variation which is unfavorable from a practical point of view. Different schemes have been proposed to reduce the peak-to-mean envelope power ratio (PMEPR) such as coding methods, clipping, reserved carriers, and probabilistic methods such as selective mapping and partial transmit sequence [1]–[7].

Unfortunately, the worst case PMEPR of multicarrier signals is rather high and is of the order of n where n is the number of subcarriers. On the other hand, the numerical evaluation of the distribution of PMEPR shows that encountering the worst case n is highly unlikely [8]–[13]. This in fact motivates the problem of finding the PMEPR distribution to quantify how severe that is. In [8], [9], by assuming that the multicarrier signal is a Gaussian process, an expression for the probability distribution of PMEPR is derived. This is a very strong assumption, and when the codewords are chosen from fixed constellations, is mathematically not valid for the joint distribution of n or more samples [14]. Recently, in [12], an upper bound for the PMEPR distribution is shown for quadrature amplitude modulation/phase-shift keying (QAM/PSK) with M^2 points and uniform distribution over the constellation points, and it is shown that the probability of encountering a PMEPR of greater than $(1+\epsilon)\log n$ is going to zero as n increases. On the other hand, in [13], using techniques different from ours, a lower bound for the distribution of PMEPR is obtained when codewords are uniformly distributed over a complex sphere. [13], however, does not perform an asymptotic analysis, which is what we do here. In this correspondence, we generalize the results to a larger class of constellations with even distribution over the constellation points, and we show a stronger result, namely, with high probability the PMEPR *behaves* like $\log n + O(\log\log n)$. In other words, encountering a PMEPR of less than $\log n + O(\log\log n)$ is also highly unlikely.

The results are based on a generalization of the well-known result of Halasz [15] for Littlewood trigonometric polynomials with equiprobable coefficients chosen independently from $\{+1, -1\}$ [10], [6], [12]. In summary, we show that, with probability approaching one, any codeword either with entries chosen independently from the symmetric QAM/PSK constellations or chosen uniformly from a complex sphere has PMEPR of $\log n + O(\log\log n)$ for a large number of subcarriers. We then use this result to determine the achievable rate of codes with given minimum distance and bounded PMEPR.

The rest of the correspondence is outlined as follows. Section II introduces the notation, multicarrier signals, and the PMEPR of a codeword. The lower and upper probability bounds for the PMEPR distribution are derived in Section III. In Section IV, we discuss the consequences of the bounds and we obtain a Varshamov–Gilbert (VG) type bound for the achievable rate of codes with bounded PMEPR and with given minimum Hamming distance.

II. Definition

The complex envelope of a multicarrier signal with n subcarriers may be represented as

$$s_C(t) = \sum_{i=1}^{n} c_i e^{j2\pi i f_0 t}, \qquad 0 \le t \le 1/f_0 \tag{1}$$

where f_0 is the subchannel spacing and $C = (c_1, \ldots, c_n)$ is the complex modulating vector with entries from a given complex constellation. The admissible modulating vectors are called codewords and the ensemble of all possible codewords constitute the code $\mathcal{C}$. For mathematical convenience, we define the normalized complex envelope of a multicarrier signal as

$$s_C(\theta) = \sum_{i=1}^{n} c_i e^{j\theta i}, \qquad 0 \le \theta < 2\pi. \tag{2}$$

Then, the PMEPR of each codeword C in the code $\mathcal{C}$ may be defined as

$$\text{PMEPR}_{\mathcal{C}}(C) = \max_{0 \le \theta < 2\pi} \frac{|s_C(\theta)|^2}{E\{\|C\|^2\}}. \tag{3}$$

Similarly, the PMEPR of the code $\mathcal{C}$, denoted by $\text{PMEPR}_{\mathcal{C}}$, is defined as the maximum of (3) over all codewords in $\mathcal{C}$. It is clear from the definition of PMEPR that if all the carriers add up coherently, the PMEPR can be of the order of n.

In this correspondence, we will consider two classes of codes, namely, complex symmetric q-ary codes in which each coordi-

Manuscript received January 6, 2003; revised October 14, 2003. This work was supported in part by the National Science Foundation under Grant CCR-0133818, by the Office of Naval Research under Grant N00014-02-1-0578, and by Caltech's Lee Center for Advanced Networking.

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: masoud@systems.caltech.edu; hassibi@systems.caltech.edu).

Communicated by G. Caire, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2004.826681