A Performance Bound on Random-Coded MIMO Systems

Jingqiao Zhang, *Student Member, IEEE*, and Heung-No Lee, *Member, IEEE*

***Abstract*— A closed-form upper bound on the average error probability is proposed for random block codes operating in multi-input multi-output (MIMO) systems. The bound exponentially decays to zero with increasing block length, and the obtained error exponent proves consistent with Gallager's random coding exponent and the information-theoretic channel capacity.**

***Index Terms*— Union bound, random coding, error exponent, MIMO system.**

I. Introduction

MIMO systems have been attracting increasing research interests since the information-theoretic channel capacity was identified by Telatar [1], and Foschini and Gans [2]. One of the interests lies in the performance evaluation of the MIMO system with discrete-alphabet inputs, since it makes much practical sense. Canonical measures, such as the information-theoretic capacity and Gallager's random coding arguments [3], can be easily extended to these discrete-input systems. Nevertheless, as stated in [4], [5], these measures generally need to be evaluated by Monte-Carlo methods because of the involved expectation/integral over the high-dimensional input signals, random fading and noises.

Similar to [3], we are concerned about the classical random coding, and propose a closed-form union bound on the average error probability of random coded MIMO systems. This bound exponentially decays to zero with increasing block length for any transmission rate less than the derived closed-form cutoff rate. It proves effective in a variety of modulation and channel scenarios by comparing with the canonical measures.

II. System of Interest

Consider an ensemble $\mathbb{C}$ of random block codes of length L and rate $R_c = K/L$. Each code in the ensemble is constructed by randomly selecting 2^K codewords out of a total number of 2^L distinct bit strings of length L, without replacement. The ensemble is composed of all $|\mathbb{C}| = \binom{2^L}{2^K}$ distinct codes generated in this manner. It is assumed that each code $\mathcal{C}$ in the ensemble is selected for use with equal probability.

As illustrated in Fig. 1, one random code is used to operate in an M-transmit N-receive MIMO system. A codeword c is uniformly chosen from the code for transmission. Each group of MK_b bits of c is modulated onto an $M \times 1$ vector $\mathbf{s}$ of symbols, whose entries take on values from a channel-symbol constellation of size 2^{K_b}. For example, $K_b = 2$ for 4PSK.

Manuscript received August 29, 2005. The associate editor coordinating the review of this letter and approving it for publication was Dr. Murat Uysal. This work was supported in part by ADCUS, Inc., Wexford, PA, and the University of Pittsburgh through the CRDF fund.

The authors are with the Electrical and Computer Engineering Department, University of Pittsburgh (e-mail: jiz33@pitt.edu, hnlee@ee.pitt.edu).

Digital Object Identifier 10.1109/LCOMM.2006.03018.

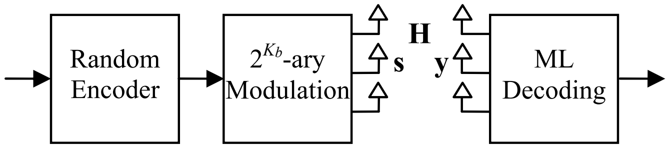

Fig. 1. Random coded MIMO systems.

The collection of all $J = 2^{MK_b}$ distinct groups of MK_b bits is denoted as $\{b_0, b_1, \cdots, b_{J-1}\}$, and that of the J corresponding symbol vectors $\mathbf{s}$ as $\{\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{J-1}\}$. We assume the symbol vector $\mathbf{s}$ obeys the average energy constraint $\mathrm{E}\|\mathbf{s}_j\|^2 = E_s$, where $\|\cdot\|$ denotes the norm of a complex vector. For convenience, assume L to be a multiple T of MK_b. The codeword c is finally transformed into an $M \times T$ space-time word $\mathbf{x} = [\mathbf{x}_1\, \mathbf{x}_2 \cdots \mathbf{x}_T]$, with $\mathbf{x}_t \in \{\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{J-1}\}$ for $t = 1, 2, \cdots, T$.

In correspondence to each symbol vector $\mathbf{s}$ in $\mathbf{x}$, an $N \times 1$ vector $\mathbf{y}$ of receive signals is obtained:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{H}$ is the $N \times M$ channel matrix whose entries are independent Rayleigh distributed random variables, and $\mathbf{n}$ is the $N \times 1$ complex, spatially and temporally white Gaussian noise with zero mean and variance N_0. The channel matrix $\mathbf{H}$ is assumed known at the receiver, and sampled independently for each symbol vector $\mathbf{s}$; i.e., the channel is ergodic.

III. Basic Definitions

In this section, we introduce a couple of definitions based on the formulation of the pairwise error probability. This will facilitate our analysis on the union upper bound later.

The pairwise error probability from codeword c to codeword c' is defined as the probability that the receiver, when performing a maximum-likelihood (ML) binary decision between the two, erroneously decides in preference of c' when c is actually transmitted. For a Rayleigh MIMO channel, the pairwise error probability of the system in (1) is formulated as [6]

$$\Pr(c \to c') \le \prod_{t=1}^{T} \left(1 + \frac{1}{4N_0}\|\mathbf{x}_t - \mathbf{x}'_t\|^2\right)^{-N}, \tag{2}$$

where $\mathbf{x}_t$ and $\mathbf{x}'_t$ are the t^{th} columns of the space-time words $\mathbf{x}$ and $\mathbf{x}'$ which are mapped from c and c', respectively.

Definition 1 (Weight Profile): Any bit string c of length L can be equivalently considered as a serial concatenation of T sub-strings of length MK_b from the set $\{b_0, b_1, \cdots, b_{J-1}\}$. Denote $\delta_j(c)$ as the number of b_j's that appear in the string c, for $j = 0, 1, \cdots, J-1$. The space-time word $\mathbf{x}$ modulated from c is thus composed of a number $\delta_j(c)$ of symbol vectors $\mathbf{s}_j$. The array $\underline{\hat{\delta}} := (\delta_0(c), \delta_1(c), \cdots, \delta_{J-1}(c))$ of these numbers is defined as the weight profile of the bit string c.

Definition 2 (Distance Profile): Consider two bit strings c and c' of length L which are modulated onto space-time words $\mathbf{x}$ and $\mathbf{x}'$, respectively. Denote $\delta_{j,k}$ as the number of the corresponding columns of $\mathbf{x}$ and $\mathbf{x}'$ that are $(\mathbf{x}_t = \mathbf{s}_j, \mathbf{x}'_t = \mathbf{s}_k)$, for $j,k = 0,1,\cdots,J-1$. The array $\underline{\delta}$ of these numbers is defined as the distance profile between c and c'. For simplicity, we denote $\underline{\delta} := (\underline{\delta}_0, \underline{\delta}_1, \cdots, \underline{\delta}_{J-1})$, where $\underline{\delta}_j := (\delta_{j,0}, \delta_{j,1}, \cdots, \delta_{j,J-1})$, for $j = 0,1,\cdots,J-1$.

It is clear that the sum of $\delta_j(c)$'s equals the total number T of the columns of one space-time word. That is,

$$\delta_j(c) \in \{0,1,\cdots,T\} \text{ and } \sum\nolimits_{j=0}^{J-1} \delta_j(c) = T. \tag{3}$$

On the other hand, since $\delta_{j,k}$ denotes the number of the combinations $(\mathbf{x}_t = \mathbf{s}_j, \mathbf{x}'_t = \mathbf{s}_k)$ and $\delta_j(c)$ counts up the columns of $\mathbf{x}$ that $\mathbf{x}_t = \mathbf{s}_j$, we have

$$\delta_{j,k} \in \{0,1,\cdots,\delta_j(c)\} \text{ and } \sum\nolimits_{k=0}^{J-1} \delta_{j,k} = \delta_j(c). \tag{4}$$

Based on the definitions above, the pairwise error probability of (2) can be concisely rewritten according to the distance profile $\underline{\delta}$ between the two codewords c and c'; i.e.,

$$\Pr(c \to c') \le \prod_{j,k=0}^{J-1} \left[\left(1 + \frac{\|\mathbf{s}_j - \mathbf{s}_k\|^2}{4N_0}\right)^{-N}\right]^{\delta_{j,k}} =: \prod_{j,k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}}, \tag{5}$$

by grouping the like terms $(\mathbf{x}_t - \mathbf{x}'_t = \mathbf{s}_j - \mathbf{s}_k)$ in (2) under each power exponent $\delta_{j,k}$. The reason for defining $\underline{\delta}$ as the distance profile is thus clear: For a given signal-to-noise ratio (SNR), $\underline{\delta}$ completely determines the upper-bound formulation of the pairwise error probability in (5).

IV. Error Performance Analysis

We next turn to the derivation of the union upper bound on the error performance. The result is summarized as follows.

Theorem: Consider the random coded MIMO system in (1). The average probability of ML decoding error over the ensemble of random codes is upper-bounded by

$$\overline{P}_e \le 2^{-T \cdot E(R)}, \tag{6}$$

where the error exponent $E(R)$ is given by

$$E(R) = \left[MK_b - \log_2\left(\frac{1}{2^{MK_b}} \sum_{j,k=0}^{2^{MK_b}-1} \beta_{j,k}\right)\right] - R =: R_0 - R, \tag{7}$$

with $R = R_c M K_b$ denoting the transmission rate of the system and R_0 serving as the cut-off rate.

Proof: The average error probability over the ensemble of random codes is upper-bounded by

$$\overline{P}_e \le \frac{1}{|\mathbb{C}|} \sum_{\mathcal{C} \in \mathbb{C}} \left\{ \frac{1}{2^K} \sum_{c \in \mathcal{C}} \left[\sum_{c' : c' \in \mathcal{C}, c' \neq c} \Pr(c \to c') \right] \right\} \tag{8}$$

$$\le \frac{1}{2^K |\mathbb{C}|} \sum_{\mathcal{C} \in \mathbb{C}} \sum_{c:\ c \in \mathcal{C};\ c' : c' \in \mathcal{C}, c' \neq c} \left[\prod_{j,k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right], \tag{9}$$

where the two outer sums in (8) are taken over the uniform choices of (i) all codes $\mathcal{C}$ from the ensemble $\mathbb{C}$, and (ii) all 2^K codewords c from the drawn code $\mathcal{C}$, respectively. Conditioned on the transmission of one codeword $c \in \mathcal{C}$, the inner sum is merely the union bound on the error performance, i.e., the sum of the pairwise error probabilities from c to any other codeword $c' \in \mathcal{C}$. Equation (9) is obtained according to (5).

It is clear the pilot codeword c in (8) and (9) should be chosen from the 2^K valid codewords within each drawn code. From the perspective of all codes in the ensemble, however, each and every 2^L distinct bit strings of length L can be considered as the pilot codeword at least once. From this point of view, the order of the sums in (9) can be changed as follows:

$$\overline{P}_e \le \frac{1}{2^K |\mathbb{C}|} \sum_{c \in GF(2)^L} \sum_{c' : c' \neq c, (c,c') \in \mathcal{C}, \mathcal{C} \in \mathbb{C}} \left[\prod_{j,k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right], \tag{10}$$

where $GF(2)^L$ denotes the set of all 2^L distinct bit strings of length L. For a given c, the inner sum is to count in all codewords $c' \neq c$ that coexist with c in the code of the ensemble.

Note that the bracketed term in (10) is completely determined by the distance profile $\underline{\delta}$. Thus, the sum over c' can be rearranged with respect to this metric:

$$\overline{P}_e \le \frac{1}{2^K |\mathbb{C}|} \sum_{c \in GF(2)^L} \sum_{\underline{\delta} : \underline{\delta} \in \Omega(c), \underline{\delta} \neq \underline{\delta}^*} S_{\underline{\delta}}(c) \left[\prod_{j,k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right], \tag{11}$$

where $\underline{\delta}^*$ denotes the unique distance profile between c and itself; i.e., $\underline{\delta} \neq \underline{\delta}^*$ is equivalent to $c' \neq c$. $\Omega(c)$ is the set of all possible distance profiles $\underline{\delta}$ associated with c. That is,

$$\Omega(c) = \left\{ \underline{\delta} \,\middle|\, \underline{\delta}_j \in \Omega_j(c), \text{ for } j = 0,1,\cdots,J-1 \right\}, \tag{12}$$

where

$$\Omega_j(c) = \left\{ \underline{\delta}_j \,\middle|\, \delta_{j,k} \in \{0,1,\cdots,\delta_j(c)\}, \sum\nolimits_{k=0}^{J-1} \delta_{j,k} = \delta_j(c) \right\}. \tag{13}$$

For each $\underline{\delta}$, there are a number of codewords c' in one code that have this distance from a given c, and $S_{\underline{\delta}}(c)$ is to add up these numbers over each code that already contains c as a codeword. $S_{\underline{\delta}}(c)$ can be calculated by a combinatorial method:

$$S_{\underline{\delta}}(c) = \left[\frac{2^K}{2^L} |\mathbb{C}|\right] \left[\frac{2^K - 1}{2^L - 1}\right] \left[\prod_{j=0}^{J-1} \binom{\delta_j(c)}{\delta_{j,0}, \cdots, \delta_{j,J-1}} \right] \text{ for } \underline{\delta} \neq \underline{\delta}^*; \tag{14}$$

where the first term is the number of codes in the ensemble that include c as a codeword. For each of these codes, the second term indicates that 2^K–1 bit strings out of 2^L (other than c) are also selected as codewords, and among all the 2^L available, the last term gives the number of the strings that have a distance $\underline{\delta}$ from c, i.e., satisfy the J constraints in (12).

Plugging (14) into (11) and ignoring the term $\underline{\delta} \neq \underline{\delta}^*$, we have

$$\overline{P}_e \le \frac{1}{2^L} \frac{2^K - 1}{2^L - 1} \sum_{c \in GF(2)^L} \sum_{\underline{\delta} \in \Omega(c)} \prod_{j=0}^{J-1} \left[\binom{\delta_j(c)}{\delta_{j,0}, \cdots, \delta_{j,J-1}} \prod_{k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right]. \tag{15}$$

Notice that for each given c, the J constraints in $\Omega(c)$ are uncorrelated. The sum over $\underline{\delta}$ in (15) can thus be simplified as

$$\sum_{\underline{\delta} \in \Omega(c)} \prod_{j=0}^{J-1} \left[\binom{\delta_j(c)}{\delta_{j,0}, \cdots, \delta_{j,J-1}} \prod_{k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right]$$

$$= \sum_{\underline{\delta}_0 \in \Omega_0(c)} \sum_{\underline{\delta}_1 \in \Omega_1(c)} \cdots \sum_{\underline{\delta}_{J-1} \in \Omega_{J-1}(c)} \prod_{j=0}^{J-1} \left[\binom{\delta_j(c)}{\delta_{j,0}, \cdots, \delta_{j,J-1}} \prod_{k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right]$$

$$= \prod_{j=0}^{J-1} \sum_{\underline{\delta}_j \in \Omega_j(c)} \left[\binom{\delta_j(c)}{\delta_{j,0}, \cdots, \delta_{j,J-1}} \prod_{k=0}^{J-1} \beta_{j,k}^{\delta_{j,k}} \right] = \prod_{j=0}^{J-1} \left(\sum_{k=0}^{J-1} \beta_{j,k} \right)^{\delta_j(c)} \tag{16}$$

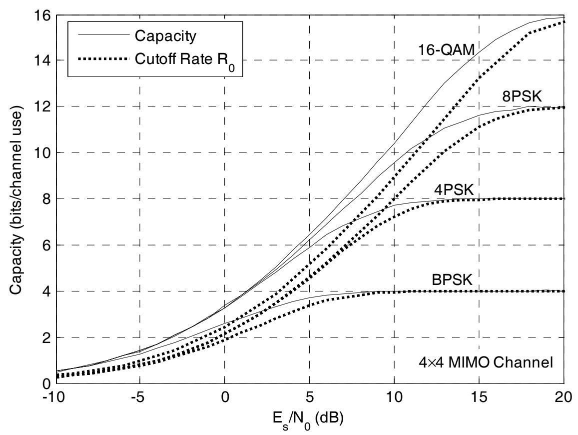

Fig. 2. Cutoff rate v.s. channel capacity (4×4 MIMO channels).

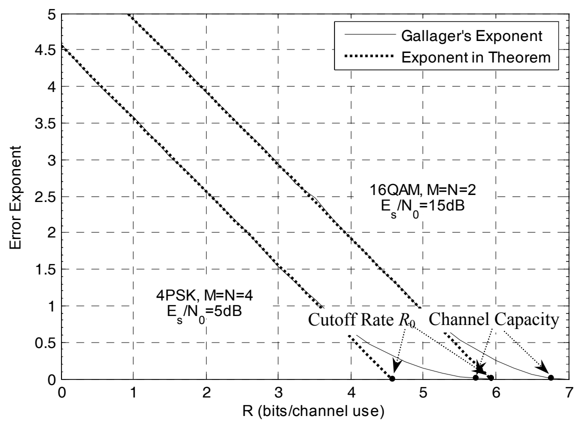

Fig. 3. A comparison between the two error exponents.

where the second equality is obtained by summing $\underline{\delta}_j$'s separately over each term in the product, and the last equality follows from the multinomial theorem in [7].

Plugging (16) into (15) and by simple manipulation, we have

$$\overline{P}_e \leq \frac{1}{2^L}\frac{2^K}{2^L}\sum_{c\in GF(2)^L}\prod_{j=0}^{J-1}\left(\sum_{k=0}^{J-1}\beta_{j,k}\right)^{\delta_j(c)}. \quad (17)$$

Similar to c', the sum over all bit strings c of length L can be reorganized with respect to their weight profiles:

$$\overline{P}_e \leq \frac{1}{2^L}\frac{2^K}{2^L}\sum_{\underline{\hat{\delta}}\in\widehat{\Omega}}\widehat{A}_{\underline{\hat{\delta}}}\prod_{j=0}^{J-1}\left(\sum_{k=0}^{J-1}\beta_{j,k}\right)^{\delta_j(c)}, \quad (18)$$

where $\widehat{\Omega}$ is the set of all possible weight profiles $\underline{\hat{\delta}}$ and according to (3) we have

$$\widehat{\Omega} = \left\{\underline{\hat{\delta}}\,\middle|\,\delta_j(c)\in\{0,\ 1,\ \cdots,\ T\},\sum\nolimits_{j=0}^{J-1}\delta_j(c)=T\right\}. \quad (19)$$

$\widehat{A}_{\underline{\hat{\delta}}}$ is the number of bit strings of weight profile $\underline{\hat{\delta}}$, i.e., the number of ways to arrange $\delta_j(c)$ sub-strings b_j ($j = 0, 1, \cdots, J-1$):

$$\widehat{A}_{\underline{\hat{\delta}}} = \binom{T}{\delta_0(c),\cdots,\delta_{J-1}(c)}. \quad (20)$$

Substituting (20) into (18), we have

$$\begin{aligned}\overline{P}_e &\leq \frac{1}{2^L}\frac{2^K}{2^L}\sum_{\underline{\hat{\delta}}\in\widehat{\Omega}}\binom{T}{\delta_0(c),\cdots,\delta_{J-1}(c)}\prod_{j=0}^{J-1}\left(\sum_{k=0}^{J-1}\beta_{j,k}\right)^{\delta_j(c)}\\ &= \frac{1}{2^L}\frac{2^K}{2^L}\left[\sum_{j=0}^{J-1}\left(\sum_{k=0}^{J-1}\beta_{j,k}\right)\right]^T. \quad (21)\end{aligned}$$

where the last step is based on the multinomial theorem [7].

Finally, since $R_c = K/L$, $T = L/MK_b$, and $J = 2^{MK_b}$, we get the result of (6) by rewriting (21) into an exponential form.

V. Discussions

The upper bound of (6) exponentially decays to zero with increasing block length L for any transmission rate leading to a positive error exponent $E(R)$, i.e., for any rate R less than the threshold R_0, which serves as a lower bound to the MIMO channel capacity. Fig. 2 illustrates the difference between the closed-form cutoff rate R_0 and the channel capacity which is calculated according to the formulation in [4, Eq. (30)]. As shown, R_0 is always $2.5 \sim 3.0$ dB away from the channel capacity, while the rate difference is at most 1.5 bits/channel-use at the moderate SNR region. These SNR and rate gaps do not vary much for different modulation and channel scenarios.

In Fig. 3, we compare the derived error exponent in (7) with Gallager's random coding exponent [3, Ch. 5]. The latter is evaluated by the Monte-Carlo method proposed in [5]. The two exponents have a good match in the straight-line region, where the parameter ρ used in [3] to optimize the exponent always equals one. The exponent (7) is thus thought to be able to serve as a closed-form evaluation of Gallager's random coding exponent with ρ = 1, although we note that the underlying reasoning to arrive at them are different. On the other hand, the two exponents diverge when the transmission rate is close to the channel capacity. The final difference between R_0 and the channel capacity is about 1 bit/channel use. This is consistent with the result in Fig. 2.

VI. Conclusion

We propose an exponential-form union bound on the average error probability for random-coded MIMO systems. The obtained error exponent leads to a closed-form expression of the cutoff rate. The comparison with the MIMO channel capacity and Gallager's random coding exponent shows that our results are effective for a variety of modulation and channel scenarios.

References

[1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] R. G. Gallager, *Information Theory and Reliable Communications*. Wiley, 1968.

[4] 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.

[5] M. S. J. Jalden and B. Ottersten, "On the random coding exponent of multiple antenna systems using space-time block codes," in *Proc. IEEE Intl. Symposium Inform. Theory*, Chicago, USA, July 2004, p. 188.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[7] R. A. Brualdi, *Introductory Combinatorics(3rd Ed.)*. Prentice Hall, 1998.