On the Relationship Between Mutual Information and Bit Error Probability for Some Linear Dispersion Codes

Xianglan Jin, Jae-Dong Yang, Kyoung-Young Song, Jong-Seon No, *Member, IEEE*, and Dong-Joon Shin, *Member, IEEE*

***Abstract*—In this paper, we derive the relationship between the bit error probability (BEP) of maximum a posteriori (MAP) bit detection and the bit minimum mean square error (BMMSE). By using this result, the relationship between the mutual information and the BEP is derived for multiple-input multiple-output (MIMO) communication systems with the bit-linear linear-dispersion (BLLD) codes for the Gaussian channel. From the relationship, the lower and upper bounds on the mutual information can be derived.**

***Index Terms*—Bit error probability (BEP), bit-linear linear-dispersion (BLLD) codes, maximum a posteriori (MAP) bit detection, minimum mean square error (MMSE), multiple-input multiple-output (MIMO), mutual information.**

I. Introduction

In the analysis of communication systems, the error probability and the minimum mean square error (MMSE) are very important performance criteria. The bit error probability (BEP) of the multiple-input multiple-output (MIMO) communication systems has been extensively studied and many results have been obtained. The mutual information can also be used for measuring the performance of communication systems and is widely studied [1]–[4]. Recently, Guo, Shamai, and Verdú [5] derived an interesting relationship between the mutual information and the MMSE for the Gaussian channel. Lozano, Tulino, and Verdú [6] obtained an approximation form of the mutual information for the single-input single-output (SISO) system with binary phase shift keying (BPSK) and quadrature phase shift keying (QPSK) in high signal to noise ratio (SNR) region. Since the relationship between the mutual information and the BEP for MIMO systems has not been found, we derive this relationship for some linear dispersion codes.

In this paper, we consider the maximum a posteriori (MAP) bit detection for MIMO systems and use BEP to denote the BEP of MAP bit detection and bit MMSE (BMMSE) to denote the MMSE in estimating an information bit for any coding and modulation schemes. Then, the relationship between the BEP and the BMMSE is derived. Using the result in [5], the relationship between the mutual information and the BEP for MIMO systems with bit-linear linear-dispersion (BLLD) codes [7] is derived in the Gaussian channel if their dispersion matrices satisfy a given condition. From the relationship, the lower and upper bounds on the mutual information can be derived by using the BEP.

The following notations will be used in this paper: capital letter denotes matrix; underscore denotes vector; boldfaced letter denotes random object; I_n denotes the $n \times n$ identity matrix; $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ mean the real and imaginary parts of a complex value, respectively; $||\cdot||$ denotes the Frobenius norm of a matrix; $\mathsf{E}\{\cdot\}$ is the expectation; the superscripts $(\cdot)^\mathsf{T}$, $(\cdot)^*$ and $(\cdot)^\dagger$ denote transpose, complex conjugation, and complex conjugate transpose, respectively; finally, $\mathrm{vec}(\cdot)$ and $\mathrm{tr}(\cdot)$ represent the vectorization and trace of a matrix.

II. BEP of MAP Detection and BMMSE

Let L_t and L_r be the numbers of transmit and receive antennas in a MIMO communication system, respectively. Let $\underline{\mathbf{x}} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{L_b}]^\mathsf{T}$ be an information vector consisting of independent binary bits $\mathbf{x}_i \in \{-1, 1\}$ and $\underline{\mathrm{f}}(\underline{\mathbf{x}})$ a bijective function corresponding to coding and modulation schemes. We assume that the average transmitted power is ρ and the perfect channel state information is available at the receiver. Then, the output signal $\underline{\mathbf{y}}$ is given as

$$\underline{\mathbf{y}} = \sqrt{\rho}\mathbf{H}\underline{\mathrm{f}}(\underline{\mathbf{x}}) + \underline{\mathbf{n}} \tag{1}$$

where $\mathbf{H}$ is an $L_r \times L_t$ channel matrix with random entries having unit power, $\underline{\mathbf{n}}$ is an $L_r \times 1$ column noise vector with random entries having unit power and being independent of $\underline{\mathbf{x}}$, and ρ represents the SNR.

MAP detection chooses $\tilde{x}_i$ to maximize the posterior probability mass function (PMF), i.e., $\tilde{x}_i = \arg\max_{x_i} P(\mathbf{x}_i = x_i | \underline{\mathbf{y}} = \underline{y})$, $i = 1, 2, \ldots, L_b$. Since we assume that x_i is a binary information bit, in this paper we will use MAP detection to denote MAP bit detection. Now, we define BMMSE which is a new performance criterion.

Definition 1: A BMMSE of $\mathbf{x}$ is the MMSE in estimating a bit $\mathbf{x}$ for a given ρ, i.e.,

$$\mathrm{bmmse}(\rho) = \mathsf{E}\{|\mathbf{x} - \hat{\mathbf{x}}(\underline{\mathbf{y}})|^2\},$$

where $\hat{\mathbf{x}}(\underline{\mathbf{y}})$ is the BMMSE estimator defined as $\hat{\mathbf{x}}(\underline{\mathbf{y}}) = \mathsf{E}\{\mathbf{x}|\underline{\mathbf{y}}\} = \sum_{x \in \{-1,1\}} xP(x|\underline{\mathbf{y}})$. □

For the MIMO system defined in (1), we have

$$\mathrm{bmmse}(\rho) = \frac{1}{L_b}\sum_{i=1}^{L_b} \mathrm{bmmse}_i(\rho),$$

Manuscript received January 9, 2007; revised May 25, 2007, October 18, 2007, and April 11, 2008; accepted April 11, 2008. The associate editor coordinating the review of this paper and approving it for publication was T. Duman.

This work was supported by the IT R&D program of MKE/IITA. [2008-F-007-01, Wireless Communication Systems in 3 Dimensional Environment].

X. Jin, J.-D. Yang, K.-Y. Song, and J.-S. No are with the Department of Electrical Engineering and Computer Science, Seoul National University, Seoul 151-744, Korea (e-mail: {xianglan.jin, yjdong, sky6174}@ccl.snu.ac.kr, jsno@snu.ac.kr).

D.-J. Shin is with the Department of Electronics and Computer Engineering, Hanyang University, Seoul 133-791, Korea (e-mail: djshin@hanyang.ac.kr).

Digital Object Identifier 10.1109/T-WC.2009.080200

where $\text{bmmse}_i(\rho) = \mathsf{E}\{|\mathbf{x}_i - \hat{\mathbf{x}}_i(\underline{\mathbf{y}})|^2\}$. Then, the relationship between the BEP and the BMMSE for the MIMO systems can be derived as follows.

Theorem 1: For the MIMO system defined in (1), the BEP of MAP detection and the BMMSE of binary information vector $\underline{\mathbf{x}}$ have the following relationships.

$$\frac{1}{4}\text{bmmse}(\rho) < P_b(\rho) < \frac{1}{2}\text{bmmse}(\rho), \quad (2a)$$

$$\lim_{\rho\to 0} P_b(\rho) = \frac{1}{2}\lim_{\rho\to 0}\text{bmmse}(\rho) \quad (2b)$$

and

$$\lim_{\rho\to \infty} P_b(\rho) = \frac{1}{4}\lim_{\rho\to \infty}\text{bmmse}(\rho). \quad (2c)$$

Proof: Let R^i_j, $j \in \{-1,1\}$, be the decision region of $\underline{\mathrm{y}}$ satisfying $P(x_i = j|\underline{\mathbf{y}} = \underline{\mathrm{y}}) > P(x_i = -j|\underline{\mathbf{y}} = \underline{\mathrm{y}})$. Then, the BEP of $\mathbf{x}_i$ can be written as

$$\begin{aligned} P_{bi} &= \sum_{x_i\in\{-1,1\}} P(x_i)P(\tilde{x}_i \neq x_i|x_i) \\ &= \int_{R^i_{-1}} p(x_i = 1, \underline{\mathrm{y}})d\underline{\mathrm{y}} + \int_{R^i_1} p(x_i = -1, \underline{\mathrm{y}})d\underline{\mathrm{y}}. \end{aligned} \quad (3)$$

The BMMSE of $\mathbf{x}_i$ can be derived as

$$\begin{aligned} &\text{bmmse}_i(\rho) \\ &= \int \frac{4p(x_i=1,\underline{\mathrm{y}})p(x_i=-1,\underline{\mathrm{y}})}{p(x_i=1,\underline{\mathrm{y}})+p(x_i=-1,\underline{\mathrm{y}})}d\underline{\mathrm{y}} \\ &= 4\int_{R^i_{-1}} \frac{p(x_i=1,\underline{\mathrm{y}})p(x_i=-1,\underline{\mathrm{y}})}{p(x_i=1,\underline{\mathrm{y}})+p(x_i=-1,\underline{\mathrm{y}})}d\underline{\mathrm{y}} \\ &+ 4\int_{R^i_{1}} \frac{p(x_i=1,\underline{\mathrm{y}})p(x_i=-1,\underline{\mathrm{y}})}{p(x_i=1,\underline{\mathrm{y}})+p(x_i=-1,\underline{\mathrm{y}})}d\underline{\mathrm{y}} \\ &= 4\int_{R^i_{-1}} \frac{p(x_i=1,\underline{\mathrm{y}})}{1+\frac{p(x_i=1,\underline{\mathrm{y}})}{p(x_i=-1,\underline{\mathrm{y}})}}d\underline{\mathrm{y}} + 4\int_{R^i_{1}} \frac{p(x_i=-1,\underline{\mathrm{y}})}{1+\frac{p(x_i=-1,\underline{\mathrm{y}})}{p(x_i=1,\underline{\mathrm{y}})}}d\underline{\mathrm{y}}. \end{aligned} \quad (4)$$

Since $0 < \frac{p(x_i=-j,\underline{\mathrm{y}})}{p(x_i=j,\underline{\mathrm{y}})} < 1$ in the region R^i_j, $j \in \{-1,1\}$, we have the following inequality

$$\begin{aligned} \frac{1}{2}\int_{R^i_j} p(x_i=-j,\underline{\mathrm{y}})d\underline{\mathrm{y}} &< \int_{R^i_j} \frac{p(x_i=-j,\underline{\mathrm{y}})}{1+\frac{p(x_i=-j,\underline{\mathrm{y}})}{p(x_i=j,\underline{\mathrm{y}})}}d\underline{\mathrm{y}} \\ &< \int_{R^i_j} p(x_i=-j,\underline{\mathrm{y}})d\underline{\mathrm{y}}. \end{aligned} \quad (5)$$

Using (3), (4), and (5), we have the inequality

$$\frac{1}{4}\text{bmmse}_i(\rho) < P_{bi}(\rho) < \frac{1}{2}\text{bmmse}_i(\rho),$$

and surely

$$\frac{1}{4}\text{bmmse}(\rho) < P_b(\rho) < \frac{1}{2}\text{bmmse}(\rho).$$

As ρ goes to 0, $\frac{p(x_i=-j,\underline{\mathrm{y}})}{p(x_i=j,\underline{\mathrm{y}})}$ approaches to 1 in the region R^i_j and we have

$$\lim_{\rho\to 0} P_{bi}(\rho) = \frac{1}{2}\lim_{\rho\to 0}\text{bmmse}_i(\rho)$$

and also, as ρ goes to infinity, $\frac{p(x_i=-j,\underline{\mathrm{y}})}{p(x_i=j,\underline{\mathrm{y}})}$ approaches to 0 in the region R^i_j and we have

$$\lim_{\rho\to \infty} P_{bi}(\rho) = \frac{1}{4}\lim_{\rho\to \infty}\text{bmmse}_i(\rho).$$

Therefore, we can have (2b) and (2c). □

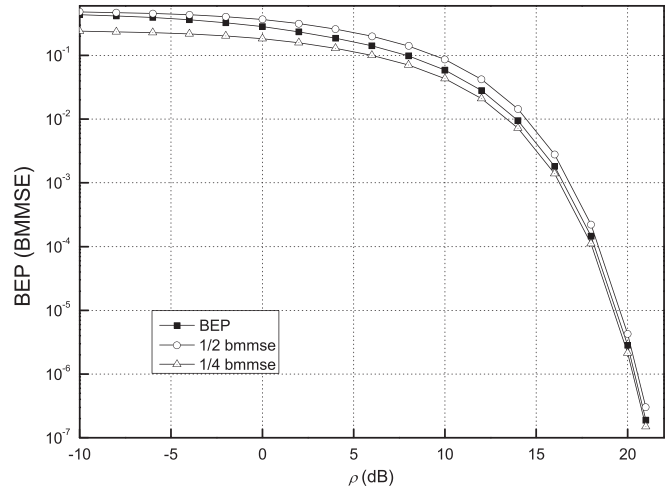

Fig. 1. The relationship between BEP and BMMSE for the SISO system with Gray coded 16QAM modulation.

An approximation similar to the proof of Theorem 1 was used to derive the relationship between the mean square error of the maximum likelihood estimator and the MMSE in [8]. As an example, we consider a SISO system $\mathbf{y} = \sqrt{\rho}f(\underline{\mathbf{x}}) + \mathbf{n}$, where $f(\cdot)$ is a Gray coded 16QAM mapper and $\mathbf{n}$ is a complex Gaussian random variable with $\mathcal{CN}(0,1)$. Using the Monte Carlo method, the BEP and the BMMSE values are plotted in Fig. 1 which shows the relationship in Theorem 1.

III. Relationship Between Mutual Information and BEP of MAP Detection for BLLD Codes

In this section, for BLLD codes, the lower and upper bounds on the mutual information are derived using the BEP. Especially, for the homogeneous orthogonal space time block codes(OSTBCs) in the Rayleigh fading channel, these bounds can be derived in closed form.

A. The Case of BLLD Codes

Let $\mathbf{X} \in \mathbb{C}^{L_t\times T}$ and $\mathbf{Y} \in \mathbb{C}^{L_r\times T}$ be the transmit and receive signal matrices, respectively, where $\mathbb{C}$ denotes the set of complex numbers and T denotes the number of symbol durations. Let $\mathbf{H} \in \mathbb{C}^{L_r\times L_t}$ be the channel matrix which is known to the receiver only. $\mathbf{H}$ does not change within T symbol durations and changes independently from block to block. Then, the MIMO system in the Gaussian channel can be expressed as

$$\mathbf{Y} = \sqrt{\rho}\mathbf{H}\mathbf{X} + \mathbf{N} \quad (6)$$

where the elements of $\mathbf{N}$ are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with mean zero and variance 0.5 per dimension, i.e., $\mathbf{N} \sim \mathcal{CN}(0, I_{L_rT})$.

A BLLD code $\mathfrak{C}$ defined in [7] is given as

$$\mathfrak{C}=\left\{X : X=\sum_{k=1}^{L_b} x_k A_k,\ x_k \in\{-1,1\},\ k=1,2,\ldots,L_b\right\} \tag{7}$$

where $A_k \in \mathbb{C}^{L_t\times T}$ are dispersion matrices and $x_1, x_2, \ldots, x_{L_b}$ are the information bits.

Let

$$\mathbf{H}'=[\mathrm{vec}(\mathbf{H}A_1),\mathrm{vec}(\mathbf{H}A_2),\ldots,\mathrm{vec}(\mathbf{H}A_{L_b})]$$

and $\underline{\mathbf{x}}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{L_b}]^{\mathsf{T}}$. Then, the MIMO system in (6) using BLLD code can be rewritten as

$$\underline{\mathbf{y}}=\sqrt{\rho}\,\mathbf{F}\underline{\mathbf{x}}+\underline{\mathbf{n}} \tag{8}$$

where $\underline{\mathbf{y}}=\begin{bmatrix}\mathrm{Re}(\mathrm{vec}(\mathbf{Y}))\\ \mathrm{Im}(\mathrm{vec}(\mathbf{Y}))\end{bmatrix}$, $\underline{\mathbf{n}}=\begin{bmatrix}\mathrm{Re}(\mathrm{vec}(\mathbf{N}))\\ \mathrm{Im}(\mathrm{vec}(\mathbf{N}))\end{bmatrix}$, and $\mathbf{F}=\begin{bmatrix}\mathrm{Re}(\mathbf{H}')\\ \mathrm{Im}(\mathbf{H}')\end{bmatrix}$. Clearly, this satisfies Theorem 1. In (8), for the fixed F, the MMSE of $F\underline{\mathbf{x}}$ is $\mathsf{E}\{||F(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}}))||^2\}$ and the mutual information $I(\rho|\mathbf{F}=F)$ of $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$ is a function of ρ. Then, the mutual information and the MMSE of $F\underline{\mathbf{x}}$ satisfy the following relationship for any input statistics [5]

$$\frac{d}{d\rho}I\big(\rho\big|\mathbf{F}=F\big)=\mathsf{E}\{||F(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}}))||^2\}\ \log_2 e. \tag{9}$$

Using Theorem 1 and (9), we derive the following relationship.

Theorem 2: Let $X=\sum_{k=1}^{L_b} x_k A_k$ be a BLLD code where L_b denotes the number of information bits during T symbol durations. Suppose that A_k's satisfy the condition

$$A_iA_j^{\dagger}+A_jA_i^{\dagger}=0,\ 1\le i<j\le L_b. \tag{10}$$

Then, for the MIMO system in (6) the relationship between the mutual information and the BEP of MAP detection of $\mathbf{x}_i$ can be derived as

$$\begin{aligned}2\log_2 e\sum_{i=1}^{L_b}||HA_i||^2P_{bi}\big(\rho\big|\mathbf{H}=H\big)&<\frac{d}{d\rho}I\big(\rho\big|\mathbf{H}=H\big)\\ &<4\log_2 e\sum_{i=1}^{L_b}||HA_i||^2P_{bi}\big(\rho\big|\mathbf{H}=H\big),\end{aligned} \tag{11a}$$

$$\begin{aligned}&\lim_{\rho\to 0}\frac{d}{d\rho}I\big(\rho\big|\mathbf{H}=H\big)\\ &=2\log_2 e\sum_{i=1}^{L_b}||HA_i||^2\lim_{\rho\to 0}P_{bi}\big(\rho\big|\mathbf{H}=H\big),\end{aligned} \tag{11b}$$

$$\begin{aligned}&\lim_{\rho\to \infty}\frac{d}{d\rho}I\big(\rho\big|\mathbf{H}=H\big)\\ &=4\log_2 e\sum_{i=1}^{L_b}||HA_i||^2\lim_{\rho\to \infty}P_{bi}\big(\rho\big|\mathbf{H}=H\big).\end{aligned} \tag{11c}$$

Proof: Using the previously defined F and $\underline{\mathbf{x}}$, the MMSE of $F\underline{\mathbf{x}}$ can be given as

$$\mathsf{E}\{\|F(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}})\|^2\}=\mathsf{E}\big\{(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}})^{\mathsf{T}}F^{\mathsf{T}}F(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}})\big\}.$$

If F satisfies

$$F^{\mathsf{T}}F=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_{L_b}) \tag{12}$$

where $\lambda_i>0$ and $\mathrm{diag}(\cdot)$ denotes the diagonal matrix, we have $\mathsf{E}\{||F(\underline{\mathbf{x}}-\hat{\underline{\mathbf{x}}})||^2\}=\sum_{i=1}^{L_b}\lambda_i\mathsf{E}\{|\mathbf{x}_i-\hat{\mathbf{x}}_i|^2\}=\sum_{i=1}^{L_b}\lambda_i\mathrm{bmmse}_i(\rho|\mathbf{H}=H)$. Then, we have the relationship between the mutual information and the BEP in (11a), (11b), and (11c). Thus, it is enough to prove that (12) holds if (10) is satisfied. The element at the jth row and ith column of $F^{\mathsf{T}}F$ is equal to $\mathrm{Re}\{\mathrm{tr}(HA_iA_j^{\dagger}H^{\dagger})\}$. Since

$$\mathrm{Re}\{\mathrm{tr}(HA_iA_j^{\dagger}H^{\dagger})\}=\frac{1}{2}\mathrm{tr}\{H(A_iA_j^{\dagger}+A_jA_i^{\dagger})H^{\dagger}\},\quad 1\le i<j\le L_b,$$

from (10), $F^{\mathsf{T}}F=\mathrm{diag}(||HA_1||^2,\ldots,||HA_{L_b}||^2)$. Thus, the theorem is proved. □

Several examples satisfying Theorem 2 are given as follows.

Example 1 (A single transmit antenna system with BPSK or QPSK): For BPSK, $A=1$ and $AA^{\dagger}=1$, and for QPSK, $A_1=1/\sqrt{2}$, $A_2=j/\sqrt{2}$, and $A_1A_2^{\dagger}+A_2A_1^{\dagger}=0$. Therefore, the dispersion matrices satisfy the condition (10) in Theorem 2.

Example 2 (Generalized linear complex OSTBCs): The generalized linear complex OSTBCs [9] can be written as

$$\left\{X : X=\sum_{i=1}^{L_b}x_iA_i, x_i\in\mathbb{R}, A_i\in\mathbb{C}^{L_t\times T},\ i=1,2,\ldots,L_b\right\}$$

and have the property

$$XX^{\dagger}=\mathrm{diag}\left(\sum_{i=1}^{L_b}l_{1,i}x_i^2,\sum_{i=1}^{L_b}l_{2,i}x_i^2,\ldots,\sum_{i=1}^{L_b}l_{L_t,i}x_i^2\right).$$

It is equivalent to

$$\begin{aligned}&A_iA_i^{\dagger}=\mathrm{diag}(l_{1,i},l_{2,i},\ldots,l_{L_t,i}),1\le i\le L_b,\\ &A_iA_j^{\dagger}+A_jA_i^{\dagger}=0,\ 1\le i<j\le L_b\end{aligned}$$

where $l_{k,i}, k=1,2,\ldots,L_t, i=1,2,\ldots,L_b$, are positive numbers determined by the type of the code. Therefore, when BPSK or QPSK is used, the codes become the BLLD codes satisfying Theorem 2.

Example 3 (Pseudo OSTBCs): Pseudo OSTBC, proposed by Jafarkhani [10], is defined as an $L_t\times T$ matrix X with entries that are linear combinations of the indeterminate variables $s_k\in S_k, k=1,2,\ldots,K$, and their conjugates such that

$$XX^{\dagger}=c(|s_1|^2+|s_2|^2+\cdots+|s_K|^2)I_{L_t},$$

where c is a constant and $S_k, k=1,2,\ldots,K$, are arbitrary subset of $\mathbb{C}$. When s_k can be described as a binary signal form, the pseudo OSTBCs satisfy Theorem 2. For example, when $s_1,s_2\in\{-1,1\}$, $s_3,s_4\in\{-j,j\}$, the following pseudo OSTBC satisfies Theorem 2.

$$X=\begin{bmatrix}s_1 & -s_2^* & -s_3^* & s_4\\ s_2 & s_1^* & -s_4^* & -s_3\\ s_3 & -s_4^* & s_1^* & -s_2\\ s_4 & s_3^* & s_2^* & s_1\end{bmatrix}.$$

To confirm Theorem 2, we compare the derivative of the mutual information and our bounds for Alamouti space-time code [11] with QPSK. We assume $\mathbf{N}\sim\mathcal{CN}(0,I_{L_rT})$ and then for the fixed $\mathbf{H}=H$, Fig. 2 shows that the inequalities in (11a) are satisfied and the lower and upper bounds are quite tight (within 0.3 dB) in the high SNR region.

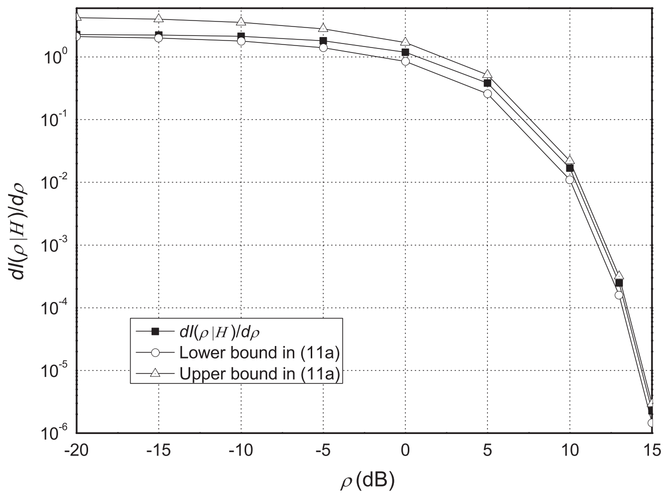

Fig. 2. The relationship between the mutual information and BEP for the Alamouti space-time code with QPSK modulation.

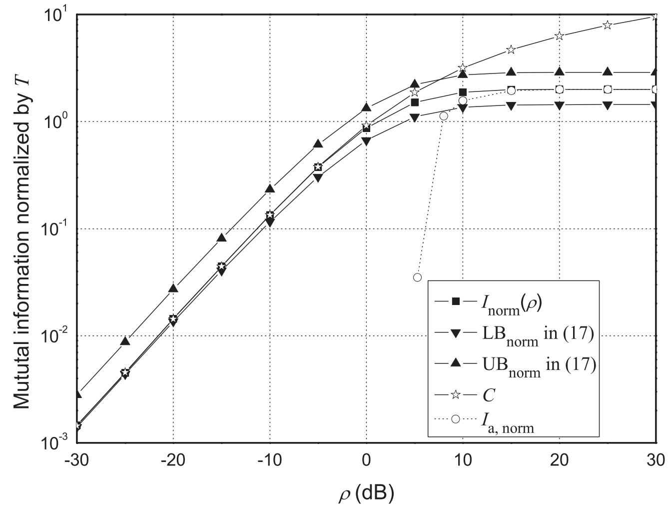

Fig. 3. Comparison of the bounds in (17), the approximation using (18), and the capacity for the Alamouti space-time code with QPSK modulation.

Integrating the terms in (11a) with respect to ρ, we obtain

$$2\log_2 e\sum_{i=1}^{L_b}||HA_i||^2\int_0^{\rho}P_{bi}\big(\gamma\big|\mathbf{H}=H\big)d\gamma < I\big(\rho\big|\mathbf{H}=H\big)$$
$$< 4\log_2 e\sum_{i=1}^{L_b}||HA_i||^2\int_0^{\rho}P_{bi}\big(\gamma\big|\mathbf{H}=H\big)d\gamma. \quad (13)$$

Theorem 2 can be applied not only to the fixed H but also to the random $\mathbf{H}$ in (13) by taking the expectation as given in the following corollary.

Corollary 1: The average mutual information of $\mathbf{X}$ and $\mathbf{Y}$ of the MIMO system in (6) with the BLLD codes satisfying (10) has the following lower and upper bounds

$$2\log_2 e\sum_{i=1}^{L_b}\mathsf{E}\Big\{||\mathbf{H}A_i||^2\int_0^{\rho}P_{bi}\big(\gamma\big|\mathbf{H}\big)d\gamma\Big\} < I(\rho)$$
$$< 4\log_2 e\sum_{i=1}^{L_b}\mathsf{E}\Big\{||\mathbf{H}A_i||^2\int_0^{\rho}P_{bi}\big(\gamma\big|\mathbf{H}\big)d\gamma\Big\}. \quad (14)$$

□

B. Calculation of Mutual Information for Homogeneous OSTBCs

In this subsection, we will simplify the bounds in (14) when the homogeneous OSTBCs defined in [12] are used with BPSK or QPSK in the Rayleigh fading channel.

The homogeneous OSTBCs satisfy the conditions

$$\begin{aligned} A_iA_i^\dagger &= cI_{L_t},\ 1\le i\le L_b,\\ A_iA_j^\dagger + A_jA_i^\dagger &= 0,\ 1\le i<j\le L_b. \end{aligned} \quad (15)$$

Thus, we can use (14) to find the bounds on the mutual information. We assume that the information bits $x_j \in \{-1,1\}, j=1,2,\ldots,L_b$, are equiprobable. From (15), we have

$$F^{\mathsf{T}}F = c\ \mathrm{tr}(HH^\dagger)I_{L_b} = c\|H\|^2 I_{L_b}$$

as in the proof of Theorem 2, and

$$\check{\underline{\mathrm{x}}} = \frac{1}{\sqrt{\rho}}(F^{\mathsf{T}}F)^{-1}F^{\mathsf{T}}\underline{\mathrm{y}} = \frac{F^{\mathsf{T}}\underline{\mathrm{y}}}{c\sqrt{\rho}\|H\|^2}$$

with $\check{\underline{\mathrm{x}}} \sim \mathcal{N}(\underline{\mathrm{x}}, \frac{1}{2c||H||^2\rho}I_{L_b})$. Then, each information bit can be detected separately and thus the BEP for the fixed H is equal to $Q(\sqrt{2c||H||^2\rho})$. Hence we obtain the lower and upper bounds on the mutual information in (14) as

$$2L_b\log_2 e\ \mathsf{E}\Big\{c\|\mathbf{H}\|^2\int_0^{\rho}Q\big(\sqrt{2c\|\mathbf{H}\|^2\gamma}\big)d\gamma\Big\} < I(\rho)$$
$$< 4L_b\log_2 e\ \mathsf{E}\Big\{c\|\mathbf{H}\|^2\int_0^{\rho}Q\big(\sqrt{2c\|\mathbf{H}\|^2\gamma}\big)d\gamma\Big\}. \quad (16)$$

Using the result in [13], we can transform the expectation value in (16) as

$$\mathsf{E}\Big\{c\|\mathbf{H}\|^2\int_0^{\rho}\frac{1}{\pi}\int_0^{\pi/2}\exp\Big(-\frac{c\|\mathbf{H}\|^2\gamma}{\sin^2\theta}\Big)d\theta d\gamma\Big\}$$
$$= \frac{c}{\pi}\int_0^{\rho}\int_0^{\pi/2}\mathsf{E}\Big\{\|\mathbf{H}\|^2\exp\Big(-\frac{c\|\mathbf{H}\|^2\gamma}{\sin^2\theta}\Big)\Big\}d\theta d\gamma.$$

Let $\mathbf{p} = \|\mathbf{H}\|^2 = \sum_{i,j}\|\mathbf{h}_{i,j}\|^2$ and $s = -c\gamma/\sin^2\theta$, where $\mathbf{h}_{ij} \sim \mathcal{CN}(0,1)$. Then, the probability density function of $\mathbf{y} = \|\mathbf{h}_{i,j}\|^2$ is $p(\mathbf{y}=y) = e^{-y},\ y>0$, and

$$\mathsf{E}\{\mathbf{p}\exp(s\mathbf{p})\} = \frac{\partial}{\partial s}\mathsf{E}\{\exp(s\mathbf{p})\}$$
$$= L_tL_r(1-s)^{-L_tL_r-1}.$$

Then, the lower and upper bounds on the mutual information in (14) are derived as

$$\frac{L_b\log_2 e}{2} - \frac{2L_b\log_2 e}{\pi}\int_0^{\pi/2}\sin^2\theta\Big(\frac{\sin^2\theta}{\sin^2\theta + c\rho}\Big)^{L_tL_r}d\theta < I(\rho)$$
$$< L_b\log_2 e - \frac{4L_b\log_2 e}{\pi}\int_0^{\pi/2}\sin^2\theta\Big(\frac{\sin^2\theta}{\sin^2\theta + c\rho}\Big)^{L_tL_r}d\theta. \quad (17)$$

For BPSK and QPSK, Lozano, Tulino, and Verdú [6] obtained the approximation of the mutual information for SISO systems in high SNR region as follows.

$$I(\rho) \approx \log_2 m - \big(e^{-\frac{d^2}{4}\rho}\log_2 e\big)\big/\big(\sqrt{\rho/\pi}\ d/2\big) \quad (18)$$

where $m = 2$ and $d = 2$ for BPSK, and $m = 4$ and $d = \sqrt{2}$ for QPSK. Since $\underline{\check{x}} \sim \mathcal{N}\big(\underline{x}, \frac{1}{2c\|H\|^2\rho} I_{L_b}\big)$, the homogeneous OSTBCs can be decoupled into several parallel SISO channels. Thus, the approximation of the mutual information for the homogeneous OSTBCs can be obtained using (18). The approximation of the mutual information given in [14] may also be used. In Fig. 3, $I_{\text{norm}}(\rho)$ denotes the mutual information normalized by $T = 2$, which is obtained by Monte Carlo simulation, C denotes the capacity, $I_{\text{a,norm}}$ denotes the approximation of the mutual information using (18) normalized by $T = 2$, and LB_{norm} and UB_{norm} denote the lower and upper bounds in (17) normalized by $T = 2$, respectively. From Fig. 3, we can see that although the approximation of the mutual information using (18) is very accurate in the high SNR region, it cannot be used in the low SNR range ($<$ 10dB). Note that the capacity can be used as an upper bound on the mutual information. Finally, although our upper bound is not tight, it can be used in the whole SNR range and, especially, the lower bound is tight in the low SNR region.

IV. Conclusion

In this paper, BMMSE is defined for the MIMO systems with any coding and modulation schemes and the relationship between the BEP and the BMMSE is derived. Using this result, for the MIMO systems with BLLD codes in the Gaussian channel, the lower and upper bounds on the mutual information are derived by using BEP when their dispersion matrices satisfy a given condition. Especially, the lower and upper bounds on the mutual information for the MIMO systems with the homogeneous OSTBCs are derived in closed form.

References

[1] E. Telatar, “Capacity of multi-antenna Gaussian channels," *AT&T Bell Labs Tech. Rep.*, June 1995.

[2] G. J. Foschini and M. J. Gans,“On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311-335, Mar. 1998.

[3] M. Dohler and H. Aghvami, “On the approximation of MIMO capacity," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 30-34, Jan. 2005.

[4] M. Kang and M.-S. Alouini, “Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112-122, Jan. 2006.

[5] D. Guo, S. Shamai (Shitz), and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1261-1282, Apr. 2005.

[6] A. Lozano, A. M. Tulino, and S. Verdú, “Optimum power allocation for parallel Gaussian channels with arbitrary input distributions,” *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3033-3051, July 2006.

[7] Y. Jiang, R. Koetter, and A. C. Singer, “On the separability of demodulation and decoding for communications over multiple-antenna block fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2709-2712, Oct. 2003.

[8] N. Chayat and S. Shamai, “Bounds on the capacity of intertransition-time-restricted binary signaling over an AWGN channel," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1992-2006, Sept. 1999.

[9] W. Su and X.-G. Xia, “On space-time block codes from complex orthogonal designs," *Wireless Pers. Commun.*, vol. 25, no. 1, pp. 1-26, Apr. 2003.

[10] H. Jafarkhani, *Space-Time Coding Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[11] S. Alamouti, “A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.

[12] S.-H. Kim, I.-S. Kang, and J.-S. No, “Exact bit error probability of orthogonal space-time block codes with quadrature amplitude modulation," in *Proc. ISIT'03*, p. 63, 2003.

[13] J. W. Craig, “A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE MILCOM'91*, pp. 571-575, 1991.

[14] S. ten Brink, G. Kramer, and A. Ashikhmin, “Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670-678, Apr. 2004.