The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

BOUNDS ON THE MUTUAL INFORMATION FOR BIT-LINEAR LINEAR-DISPERSION CODES

Xianglan Jin, Jae-Dong Yang, Kyoung-Young Song
Department of EECS, INMC
Seoul National University
Seoul, Korea
Email: {xianglan.jin, yjdong, sky6174}@ccl.snu.ac.kr

Jong-Seon No
Department of EECS, INMC
Seoul National University
Seoul, Korea
Email: jsno@snu.ac.kr

Dong-Joon Shin
Division of ECE
Hanyang University
Seoul, Korea
Email: djshin@hanyang.ac.kr

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 (MMSE), that is, the BEP is greater than a quarter of the bit MMSE and less than a half of the bit MMSE. By using this result, the lower and upper bounds of the derivative of the mutual information are derived from the BEP in the multiple-input multiple-output (MIMO) communication systems with the bit-linear linear-dispersion (BLLD) codes in the Gaussian channel.

I. INTRODUCTION

In the analysis of communication systems, the error probability and the minimum mean square error (MMSE) are very important performance criteria. Especially, in most applications of digital wireless communication systems, their error performances are measured in bits and thus the bit error probability (BEP) and the MMSE of information bits become more significant. The BEP of 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. In general, it is very difficult to derive the mutual information of MIMO systems. Telatar [1], Foschini and Gans [2] derived the capacity (that is, the maximum mutual information) of MIMO systems in the Gaussian channel. Recently, Guo, Shamai, and Verdú [3] derived an interesting relationship between the mutual information and the MMSE in the Gaussian channel.

In this paper, we use the BEP to denote the BEP of maximum a posteriori (MAP) bit detection and the bit MMSE to denote the MMSE in estimating a bit for any coding and modulation scheme. Then, the relationship between the BEP and the bit MMSE is derived such that the BEP is between a quarter of the bit MMSE and a half of the bit MMSE.

For bit-linear linear-dispersion (BLLD) codes, there exists a linear relation between the information bits and the output codeword as shown in [4]. Using the result in [3], the lower and upper bounds of the derivative of the mutual information in the MIMO systems with the BLLD codes are derived from the BEP.

The following notations will be used in this paper: capital letters denote matrices; underscores denote vectors; boldfaced letters denote random objects; $(\cdot)^\mathsf{T}$ denotes transpose; $(\cdot)^\dagger$ denotes complex conjugate transpose; $(\cdot)^*$ denotes complex conjugation; $\|\cdot\|$ denotes the Frobenius norm of a matrix; vec$(\cdot)$ means the vectorization of a matrix into a column vector; tr$(\cdot)$ represents the trace of a matrix; Re$(\cdot)$ and Im$(\cdot)$ mean the real and imaginary parts of a complex value, respectively; $\mathsf{E}\{\cdot\}$ is the expectation taken over the joint distribution of the random variables.

This paper is organized as follows. The relationship between the BEP and the bit MMSE is derived in Section II. In Section III, the lower and upper bounds of the derivative of the mutual information are derived from the BEP in the MIMO systems with the BLLD codes. Finally, the conclusion is given in Section IV.

II. BEP OF MAP BIT DETECTION AND BIT MMSE

Let L_t and L_r be the numbers of transmit antennas 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 $L_b \times 1$ information vector consisting of independent binary bits $\mathbf{x}_i \in \{-1, 1\}$ and $\underline{f}(\underline{\mathbf{x}})$ a bijective function having $L_t \times 1$ column vectors as its function values. In this section, we assume that the perfect channel state information is available at the receiver. Then the output signal of MIMO communication system can be given as

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

where H is a deterministic $L_r \times L_t$ channel matrix, $\underline{\mathbf{n}}$ an $L_r \times 1$ column vector with random entries having nonzero finite power and being independent of the information vector $\underline{\mathbf{x}}$, and ρ represents the signal-to-noise ratio (SNR).

MAP detection chooses $\tilde{x}_i$ to maximize the posterior probability mass function (PMF), i.e., $\tilde{x}_i = \arg\max_{x_i} P(\mathbf{x}_i = x_i | \underline{\mathbf{y}} = \underline{y}),\ i = 1, 2, \ldots, L_b$. In this paper, we assume that x_i is a binary signal and MAP bit detection is used. Then, the BEP can be represented as

$$P_b(\rho) = \frac{1}{L_b} \sum_{i=1}^{L_b} \sum_{x_i \in \{-1,1\}} P(x_i) P(\tilde{x}_i \neq x_i | x_i; \rho). \quad (2)$$

Now, we define bit MMSE which can be a new performance criterion.

Definition 1 *A bit MMSE is the MMSE in estimating a bit $\mathbf{x}_i$, i.e.,*

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

*This work was supported by the ITRC program of the MIC, Korea, and by the MOE, the MOCIE, and the MOLAB, Korea, through the fostering project of the Lab of Excellency.

where $\hat{\mathbf{x}}_i(\underline{\mathbf{y}};\rho)$ *is the bit MMSE estimator defined as*

$$\hat{\mathbf{x}}_i(\underline{\mathbf{y}};\rho) = \mathsf{E}\{\mathbf{x}_i|\underline{\mathbf{y}};\rho\} = \sum_{x_i\in\{-1,1\}} x_i P(x_i|\underline{\mathbf{y}};\rho).$$

□

In the MIMO system with multiple binary inputs defined in (1), the bit MMSE is represented as

$$\text{bmmse}(\rho) = \frac{1}{L_b}\mathsf{E}\{\|\underline{\mathbf{x}} - \hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}};\rho)\|^2\}$$

where $\hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}};\rho) = [\hat{\mathbf{x}}_1(\underline{\mathbf{y}};\rho), \hat{\mathbf{x}}_2(\underline{\mathbf{y}};\rho), \ldots, \hat{\mathbf{x}}_{L_b}(\underline{\mathbf{y}};\rho)]^{\mathsf{T}}$. Then, the relationship between the BEP and the bit MMSE for the MIMO systems can be derived as follows.

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

$$\frac{1}{4}\text{bmmse}(\rho) < P_b(\rho) < \frac{1}{2}\text{bmmse}(\rho) \qquad \text{(3a)}$$

and

$$\lim_{\rho\to 0} P_b(\rho) = \frac{1}{2}\lim_{\rho\to 0}\text{bmmse}(\rho), \qquad \text{(3b)}$$

$$\lim_{\rho\to \infty} P_b(\rho) = \frac{1}{4}\lim_{\rho\to \infty}\text{bmmse}(\rho). \qquad \text{(3c)}$$

Proof: Let R_j^i, $j\in\{-1,1\}$, be the decision region of $\underline{\text{y}}$ satisfying $P(x_i = j|\underline{\mathbf{y}} = \underline{\text{y}};\rho) > P(x_i = -j|\underline{\mathbf{y}} = \underline{\text{y}};\rho)$. Then the BEP in (2) can be rewritten as

$$\begin{aligned} P_b(\rho) &= \frac{1}{L_b}\sum_{i=1}^{L_b}[P(x_i=1)P(\underline{\mathbf{y}}\in R_{-1}^i|x_i=1;\rho) \\ &\quad + P(x_i=-1)P(\underline{\mathbf{y}}\in R_1^i|x_i=-1;\rho)] \\ &= \frac{1}{L_b}\sum_{i=1}^{L_b}\left[\int_{R_{-1}^i} p(x_i=1,\underline{\text{y}};\rho)d\underline{\text{y}} \right. \\ &\quad \left. + \int_{R_1^i} p(x_i=-1,\underline{\text{y}};\rho)d\underline{\text{y}}\right]. \end{aligned} \qquad (4)$$

Since the bit MMSE estimator is given as

$$\begin{aligned} &\hat{\mathbf{x}}_i(\underline{\mathbf{y}} = \underline{\text{y}};\rho) \\ &= \mathsf{E}\{\mathbf{x}_i|\underline{\mathbf{y}} = \underline{\text{y}};\rho\} \\ &= \frac{p(x_i=1,\underline{\text{y}};\rho) - p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho) + p(x_i=-1,\underline{\text{y}};\rho)},\ i = 1,2,\ldots,L_b, \end{aligned}$$

the bit MMSE can be derived as

$$\begin{aligned} \text{bmmse}(\rho) &= \frac{1}{L_b}\sum_{i=1}^{L_b}\int\Big[p(x_i=1,\underline{\text{y}};\rho) \\ &\quad \cdot\Big(1 - \frac{p(x_i=1,\underline{\text{y}};\rho) - p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho) + p(x_i=-1,\underline{\text{y}};\rho)}\Big)^2 \\ &\quad + p(x_i=-1,\underline{\text{y}};\rho) \\ &\quad \cdot\Big(-1 - \frac{p(x_i=1,\underline{\text{y}};\rho) - p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho) + p(x_i=-1,\underline{\text{y}};\rho)}\Big)^2\Big]d\underline{\text{y}} \\ &= \frac{1}{L_b}\sum_{i=1}^{L_b}\int \frac{4p(x_i=1,\underline{\text{y}};\rho)p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho)+p(x_i=-1,\underline{\text{y}};\rho)}d\underline{\text{y}} \\ &= \frac{4}{L_b}\sum_{i=1}^{L_b}\Big[\int_{R_{-1}^i} \frac{p(x_i=1,\underline{\text{y}};\rho)p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho)+p(x_i=-1,\underline{\text{y}};\rho)}d\underline{\text{y}} \\ &\quad + \int_{R_1^i} \frac{p(x_i=1,\underline{\text{y}};\rho)p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho)+p(x_i=-1,\underline{\text{y}};\rho)}d\underline{\text{y}}\Big] \\ &= \frac{4}{L_b}\sum_{i=1}^{L_b}\Big[\int_{R_{-1}^i} \frac{p(x_i=1,\underline{\text{y}};\rho)}{1+\frac{p(x_i=1,\underline{\text{y}};\rho)}{p(x_i=-1,\underline{\text{y}};\rho)}}d\underline{\text{y}} \\ &\quad + \int_{R_1^i} \frac{p(x_i=-1,\underline{\text{y}};\rho)}{1+\frac{p(x_i=-1,\underline{\text{y}};\rho)}{p(x_i=1,\underline{\text{y}};\rho)}}d\underline{\text{y}}\Big]. \end{aligned} \qquad (5)$$

Since $0 < \frac{p(x_i=-j,\underline{\text{y}};\rho)}{p(x_i=j,\underline{\text{y}};\rho)} < 1$ in the region R_j^i, $j\in\{-1,1\}$, we have the following inequality

$$\begin{aligned} \frac{1}{2}\int_{R_j^i} p(x_i=-j,\underline{\text{y}};\rho)d\underline{\text{y}} &< \int_{R_j^i}\frac{p(x_i=-j,\underline{\text{y}};\rho)}{1+\frac{p(x_i=-j,\underline{\text{y}};\rho)}{p(x_i=j,\underline{\text{y}};\rho)}}d\underline{\text{y}} \\ &< \int_{R_j^i} p(x_i=-j,\underline{\text{y}};\rho)d\underline{\text{y}}. \end{aligned} \qquad (6)$$

Using (4), (5), and (6), we have the inequality

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

As ρ goes to 0, $\frac{p(x_i=-j,\underline{\text{y}};\rho)}{p(x_i=j,\underline{\text{y}};\rho)}$ approaches to 1 in the region R_j^i and we have

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

Also, as ρ goes to infinity, $\frac{p(x_i=-j,\underline{\text{y}};\rho)}{p(x_i=j,\underline{\text{y}};\rho)}$ approaches to 0 in the region R_j^i and we have

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

□

It is easy to see that for $\rho = 0$, the BEP is equal to a half of the bit MMSE since $P_b(0) = 1/2$ and bmmse(0) $= 1$. From Theorem 1, the upper and lower bounds of the BEP (or bit MMSE) can be derived by the bit MMSE (or BEP). The similar approximation as in (6) was used to derive the relationship between the mean square error of the maximum likelihood estimator and the MMSE [5]. There are two simple examples showing the relationship in Theorem 1. First, a single-input single-output (SISO) system $\mathbf{y} = \sqrt{\rho}\mathbf{x} + \mathbf{n}$ with equally probable binary information bit in the Gaussian channel $\mathbf{n}\sim\mathcal{N}(0,1)$ with zero mean and unit variance is considered. The bit MMSE is equal to the MMSE as given in [3], i.e.,

$$\text{bmmse}(\rho) = 1 - \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\tanh(\rho - \sqrt{\rho}y)dy$$

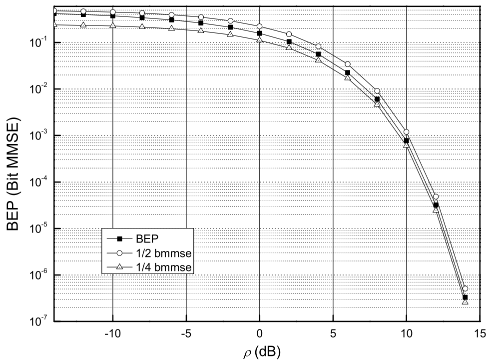

Fig. 1. The relationship between the BEP and the bit MMSE for the SISO system with binary input in the Gaussian channel.

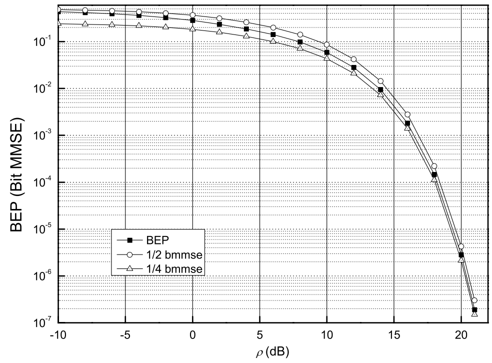

Fig. 2. The relationship between the BEP and the bit MMSE for the SISO system with Gray coded 16QAM modulation in the Gaussian channel.

and the BEP is given as $P_b = Q(\sqrt{\rho})$, where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{y^2}{2}}dy$. Fig. 1 shows the relationship of the BEP and the bit MMSE for this SISO system, which confirms the inequality in Theorem 1.

Next, we consider a SISO system $\mathbf{y} = \sqrt{\rho}g(\underline{\mathbf{x}}) + \mathbf{n}$, where $g(\cdot)$ is a Gray coded 16QAM mapper which is a complex bijective function of the four information bits, and $\mathbf{n}$ is a complex Gaussian random variable with $\mathcal{CN}(0,1)$. We assume that the average transmitted power is equal to 1. Using the Monte Carlo method, the BEP and the bit MMSE values are plotted in Fig. 2, which tells us that the relationship in Theorem 1 is also satisfied.

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

There have been many research results on the BEP of MIMO systems, but it is very difficult to derive the mutual information. In this section, the lower and upper bounds of the mutual information for the MIMO systems are derived using the BEP when the BLLD codes are used.

We consider the MIMO system with L_t transmit antennas and L_r receive antennas. 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 T denotes the number of symbol durations in each signal matrix. Let $\mathbf{H} \in \mathbb{C}^{L_r \times L_t}$ be the channel matrix, which is known to the receiver only, does not change within the interval of T symbol durations, and changes independently from block to block.

Then the MIMO system in the Gaussian channel can be expressed as

$$\mathbf{Y} = \sqrt{\rho}\,\mathbf{H}\mathbf{X} + \mathbf{N} \tag{7}$$

where the elements of $\mathbf{N}$ are independent and identically distributed (i.i.d.) and circularly symmetric complex Gaussian random variables with mean zero and variance 0.5 per dimension and the elements $\mathbf{h}_{i,j}$ of $\mathbf{H} = [\mathbf{h}_{i,j}]$ are also i.i.d. and circularly symmetric complex Gaussian random variables with mean zero and variance 0.5 per dimension.

A BLLD code $\mathcal{C}$ defined in [4] is given as

$$\mathcal{C} = \left\{ X : X = \sum_{k=1}^{L_b} x_k A_k,\ x_k \in \{-1,1\},\ k = 1,2,\ldots,L_b \right\} \tag{8}$$

where $A_k \in \mathbb{C}^{L_t \times T}$, $1 \le k \le L_b$, are dispersion matrices and the information bits $x_1, x_2, \ldots, x_{L_b}$ are equiprobable.

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 (7) can be rewritten as

$$\mathrm{vec}(\mathbf{Y}) = \sqrt{\rho}\mathbf{H}'\underline{\mathbf{x}} + \mathrm{vec}(\mathbf{N}).$$

Further, let $\underline{\mathbf{y}} = \begin{bmatrix} \mathrm{Re}(\mathrm{vec}(\mathbf{Y})) \\ \mathrm{Im}(\mathrm{vec}(\mathbf{Y})) \end{bmatrix}$, $\underline{\mathbf{n}} = \begin{bmatrix} \mathrm{Re}(\mathrm{vec}(\mathbf{N})) \\ \mathrm{Im}(\mathrm{vec}(\mathbf{N})) \end{bmatrix}$, and $\mathbf{F} = \begin{bmatrix} \mathrm{Re}(\mathbf{H}') \\ \mathrm{Im}(\mathbf{H}') \end{bmatrix}$. Then the MIMO system can be represented as

$$\underline{\mathbf{y}} = \sqrt{\rho}\,\mathbf{F}\underline{\mathbf{x}} + \underline{\mathbf{n}}. \tag{9}$$

Clearly, this equivalent system model satisfies the relationship between the BEP and the bit MMSE in Theorem 1 as

$$\frac{1}{4}\mathrm{bmmse}\big(\rho\big|\mathbf{F}=F\big) < P_b\big(\rho\big|\mathbf{F}=F\big) < \frac{1}{2}\mathrm{bmmse}\big(\rho\big|\mathbf{F}=F\big) \tag{10a}$$

and

$$\lim_{\rho\to 0} P_b\big(\rho\big|\mathbf{F}=F\big) = \frac{1}{2}\lim_{\rho\to 0}\mathrm{bmmse}\big(\rho\big|\mathbf{F}=F\big), \tag{10b}$$

$$\lim_{\rho\to \infty} P_b\big(\rho\big|\mathbf{F}=F\big) = \frac{1}{4}\lim_{\rho\to \infty}\mathrm{bmmse}\big(\rho\big|\mathbf{F}=F\big). \tag{10c}$$

The MMSE in estimating $F\underline{\mathbf{x}}$ is $\mathsf{E}\{\|F(\underline{\mathbf{x}} - \hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}};\rho))\|^2\}$ and the mutual information of $\underline{\mathbf{x}}$ and $\underline{\mathbf{y}}$ with the deterministic F in (9) is a function of ρ given as

$$I\big(\underline{\mathbf{x}}; \sqrt{\rho}\mathbf{F}\underline{\mathbf{x}} + \underline{\mathbf{n}}\big|\mathbf{F}=F\big) = I\big(\rho\big|\mathbf{F}=F\big).$$

Then the mutual information and the MMSE of $F\underline{\mathbf{x}}$ satisfy the following relationship regardless of the input statistics [3]

$$\frac{d}{d\rho}I\big(\rho\big|\mathbf{F}=F\big) = \mathsf{E}\Big\{\Big\|F\Big(\underline{\mathbf{x}} - \hat{\underline{\mathbf{x}}}(\underline{\mathbf{y}};\rho)\Big)\Big\|^2\Big\}\ \log_2 e \tag{11}$$

where the factor $1/2$ in front of MMSE of the relation in [3] is removed because each element of the noise vector $\underline{\mathbf{n}}$ is Gaussian random variable with $\mathcal{N}(0, \frac{1}{2})$. Using (10) and (11), the relationship between the mutual information and the BEP for MIMO systems using BLLD code can be derived as in the following theorem.

Theorem 2 *Let $X = \sum_{k=1}^{L_b} x_k A_k$ be a BLLD code with dispersion matrices $A_k \in \mathbb{C}^{L_t \times T}$, $1 \le k \le L_b$, and L_b the number of transmitted information bits during T symbol durations. Suppose that the dispersion matrices A_k satisfy the conditions*

$$A_i A_i^\dagger = C,\ 1 \le i \le L_b \tag{12a}$$

and

$$A_i A_j^\dagger + A_j A_i^\dagger = 0,\ 1 \le i < j \le L_b \tag{12b}$$

where C is a constant positive semidefinite matrix. Then, for the MIMO system in (7) with the deterministic matrix $\mathbf{H} = H$, we can derive the relationship between the mutual information of $\mathbf{X}$ and $\mathbf{Y}$ and the BEP of the MAP bit detection as follows.

$$\begin{aligned} 2L_b \log_2 e \cdot \mathrm{tr}(HCH^\dagger) P_b(\rho|\mathbf{H} = H) &< \frac{d}{d\rho} I(\rho|\mathbf{H} = H) \\ &< 4L_b \log_2 e \cdot \mathrm{tr}(HCH^\dagger) P_b(\rho|\mathbf{H} = H) \end{aligned} \tag{13a}$$

and

$$\lim_{\rho \to 0} \frac{d}{d\rho} I(\rho|\mathbf{H} = H) = 2L_b \log_2 e \cdot \mathrm{tr}(HCH^\dagger) \lim_{\rho \to 0} P_b(\rho|\mathbf{H} = H), \tag{13b}$$

$$\lim_{\rho \to \infty} \frac{d}{d\rho} I(\rho|\mathbf{H} = H) = 4L_b \log_2 e \cdot \mathrm{tr}(HCH^\dagger) \lim_{\rho \to \infty} P_b(\rho|\mathbf{H} = H). \tag{13c}$$

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

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

If F satisfies

$$F^\mathsf{T} F = \lambda \mathrm{I} \tag{14}$$

where $\lambda > 0$ and I denotes the identity matrix, we have $\mathsf{E}\{\|F(\underline{\mathbf{x}} - \hat{\underline{\mathbf{x}}})\|^2\} = \lambda \mathsf{E}\{\|\underline{\mathbf{x}} - \hat{\underline{\mathbf{x}}}\|^2\} = \lambda L_b \cdot \mathrm{bmmse}(\rho)$. Then, from (10a), (10b), (10c), and (11), we get the relationship between the mutual information and the BEP in (13a), (13b), and (13c). Thus, all we have to do is to prove that (14) holds if (12a) and (12b) are satisfied. Since $F^\mathsf{T} F = Re(H'^\dagger H')$, we will consider $H'^\dagger H'$ as follows.

$$\begin{aligned} &H'^\dagger H' \\ &= [\mathrm{vec}(HA_1), \cdots, \mathrm{vec}(HA_{L_b})]^\dagger \cdot [\mathrm{vec}(HA_1), \cdots, \mathrm{vec}(HA_{L_b})] \\ &= \begin{bmatrix} (\mathrm{vec}(HA_1))^\dagger \mathrm{vec}(HA_1) & \cdots & (\mathrm{vec}(HA_1))^\dagger \mathrm{vec}(HA_{L_b}) \\ \vdots & \ddots & \vdots \\ (\mathrm{vec}(HA_{L_b}))^\dagger \mathrm{vec}(HA_1) & \cdots & (\mathrm{vec}(HA_{L_b}))^\dagger \mathrm{vec}(HA_{L_b}) \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{tr}(HA_1 A_1^\dagger H^\dagger) & \cdots & \mathrm{tr}(HA_{L_b} A_1^\dagger H^\dagger) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}(HA_1 A_{L_b}^\dagger H^\dagger) & \cdots & \mathrm{tr}(HA_{L_b} A_{L_b}^\dagger H^\dagger) \end{bmatrix}. \end{aligned}$$

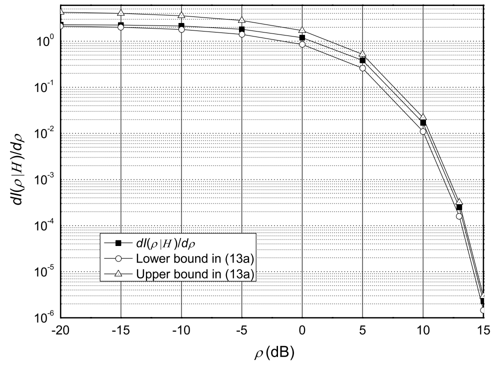

Fig. 3. The relationship between the derivative of the mutual information and the BEP for Alamouti space-time code with QPSK modulation.

Then, each component in this matrix has the following property

$$\begin{aligned} \mathrm{tr}(HA_i A_j^\dagger H^\dagger) &= \mathrm{tr}((HA_i A_j^\dagger H^\dagger)^\mathsf{T}) = \mathrm{tr}(H^* A_j^* A_i^\mathsf{T} H^\mathsf{T}) \\ &= [\mathrm{tr}(HA_j A_i^\dagger H^\dagger)]^*,\ 1 \le i < j \le L_b \end{aligned}$$

and then

$$\begin{aligned} &\mathrm{Re}\{\mathrm{tr}(HA_i A_j^\dagger H^\dagger)\} \\ &= \frac{1}{2}\{\mathrm{tr}(HA_i A_j^\dagger H^\dagger) + \mathrm{tr}(HA_j A_i^\dagger H^\dagger)\} \\ &= \frac{1}{2}\mathrm{tr}\{H(A_i A_j^\dagger + A_j A_i^\dagger)H^\dagger\},\ 1 \le i < j \le L_b. \end{aligned}$$

From (12a) and (12b), $F^\mathsf{T} F = \mathrm{Re}(H'^\dagger H') = \mathrm{tr}(HCH^\dagger) \cdot \mathrm{I}$. Thus, this theorem is proved. □

Two examples confirming Theorem 2 are given as follows. First, we consider a single-transmit antenna system using BPSK or QPSK. For BPSK case, $A = 1$ and $AA^\dagger = 1$, and for QPSK case, $A_1 = 1/\sqrt{2}$, $A_2 = j/\sqrt{2}$ and $A_1 A_1^\dagger = 1/2$, $A_2 A_2^\dagger = 1/2$, $A_1 A_2^\dagger + A_2 A_1^\dagger = 0$. Therefore, the dispersion matrices satisfy the conditions (12a) and (12b) in Theorem 2.

Second, we consider the homogeneous orthogonal space-time block codes (OSTBCs) $\mathcal{L}$ defined by Kim, Kang, and No [6], which can be rewritten as

$$\mathcal{L} = \left\{ X : X = \sum_{k=1}^{L_b} x_k A_k, x_k \in \mathbb{R}, A_k \in \mathbb{C}^{L_t \times T},\ k = 1, 2, \ldots, L_b \right\}$$

and have the property $XX^\dagger = c \sum_{k=1}^{L_b} x_i^2 \cdot \mathrm{I}$. From the property of OSTBCs in [7, II], it can be equivalent to

$$\begin{aligned} A_i A_i^\dagger &= c\mathrm{I},\ 1 \le i \le L_b, \\ A_i A_j^\dagger + A_j A_i^\dagger &= 0,\ 1 \le i < j \le L_b \end{aligned}$$

where c is a constant determined by the type of homogeneous OSTBC. When BPSK or QPSK is used, the homogeneous OSTBCs become the BLLD codes and the conditions (12a) and (12b) in Theorem 2 are satisfied. Then we conclude that for BPSK or QPSK, the homogenous OSTBCs satisfy the inequalities (13a), (13b) and (13c) in Theorem 2. Fig. 3 shows the three terms in (13a) for the Alamouti space-time code [8] with QPSK symbols, from which we conclude that the inequalities in (13a) are satisfied and the lower and upper bounds are quite tight in the low and high SNR regions, respectively.

IV. Conclusion

In this paper, bit MMSE is defined for the MIMO systems with any coding and modulation schemes and the relationship between the BEP and the bit MMSE is derived. Using this result, the lower and upper bounds of the derivative of the mutual information for the MIMO systems with BLLD codes are derived in terms of BEP if their dispersion matrices satisfy some conditions.

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 Personal Communications*, vol. 6, pp. 311–335, Mar. 1998

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

[4] 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. Inf. Theory*, vol. 49, no. 10, pp. 2709–2712, Oct. 2003.

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

[6] 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.

[7] X. B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2468–2503, Oct. 2003.

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