# A Tight BER Upper Bound for Bit-Interleaved Coded Modulation with Square QAM and Gray Labeling

Liang-Wei Huang, Wern-Ho Sheen, *Member, IEEE,* and Tsang-Wei Yu

***Abstract*— Bit-interleaved coded modulation (BICM) is a bandwidth-efficient coded system with *diversity order* higher than that of Ungerboeck's trellis-coded modulation on fading channels. In this paper, we investigate the BER (bit error rate) performance of BICM in the additive white Gaussian and Rayleigh fading channels. A new upper bound is given for the square QAM constellation with gray labeling, which constitutes a large portion of practical applications of BICM systems. The new upper bound is tighter than the well-known BICM union bound proposed in [4].**

***Index Terms*— BER upper bounds, bit-interleaved coded modulation (BICM), QAM with gray labeling, Rayleigh fading channels.**

## I. Introduction

TRELLIS-CODED modulation (TCM) was proposed by Ungerboeck for bandwidth-efficient communications in the additive white Gaussian noise (AWGN) channels [1]. Through a joint design of coding and high-order modulation, one can achieve a sizable coding gain with no increase in the signal bandwidth. Originally, TCM was designed to maximize the minimum Euclidean distance of the coded system in order to achieve high coding gain. For fading channels, however, Divsalar and Simon showed that it is the diversity order rather than the minimum Euclidean distance that plays a primary role to achieve such an objective [2]. Using no parallel transitions and/or increasing the constraint length of the constitute code are the conventional techniques to increase the diversity order [2]. Nevertheless, they may not be as effective due to the fact that the diversity order is limited to the minimum number of distinct coded symbols along an error event. Zehavi [3] proposed to add a bit inter-leaver between the channel encoder and modulator so that the diversity order can be increased to the minimum number of distinct coded bits rather than the distinct coded symbols. This technique was later named bit-interleaved coded modulation (BICM) and shown to outperform TCM on fading channels [3], [4].

After Zehavi's pioneering work in [3], Caire, Taricco, and Biglieri laid a theoretical foundation for BICM from a viewpoint of information theory [4]. Along with new performance analysis, design guidelines were given for searching for good BICM systems. In [5], motivated by the principle of turbo decoding, BICM with iterative decoding was proposed. By using a proper labeling, the minimum inter-signal Euclidean distance can be increased without compromising the diversity order and that leads to a better performance in the AWGN channels. Furthermore, BICM was applied to multiple input multiple output (MIMO) systems in order to improve performance and/or increase data rate [6].

Most of the previous works on BER (Bit Error Rate) analysis of BICM systems were based on the BICM union and expurgated bounds proposed in [4]. The former is quite loose in general, and the latter is only an approximation theoretically except the case with QPSK and gray labeling. In this paper, a new tight upper bound is derived for the square QAM constellation with gray labeling, which constitutes a large portion of practical applications of BICM systems. Numerical results show that the new upper bound is tighter than the well-known BICM union bound in [4].

## II. System Model

The equivalent parallel, independent channel model in [4] is adopted here. Fig. 1 is the block diagram, where a constellation, $\Upsilon$, with $M = 2^m$ signal points is assumed. Each channel corresponds to one of the $m$ label positions of a signal point $x$ in $\Upsilon$. The code is a (punctured) convoultional code, although some other codes are also possible [4]. Define $\Upsilon_b^i$ be the constellation subset used by $i$-th the channel ( $i$-th label position) with the code bit $c_k = b$ , $b = 0, 1$ (see Fig. 2). Under the assumption of ideal inter-leaver, one of the channels, says the $i^{th}$ channel, is randomly selected by $c_k$, and a signal point $x_k \in \Upsilon_b^i$ is transmitted, through the operation of signal mapping $u$, over the channel. Because of the assumption of ideal interleaving, the signals in $\Upsilon_b^i$ are selected equally probably.

From Fig. 1, $y_k^i = \rho_k^i x_k^i + n_k^i$. Both the AWGN and Rayleigh fading channels with coherent demodulation are considered. For the former, $\rho_k^i = 1$ , and for the latter $\rho_k^i$ is the Rayleigh fading gain, while $n_k^i$ is a complex-valued, zero-mean Gaussian noise. With an ideal inter-leaver, $\{\rho_k^i\}$ and $\{n_k^i\}$ are i.i.d. random variables, respectively and are independent of other random variables. It was shown in [4] that a short inter-leaver would be sufficient to achieve nearly the performance of ideal inter-leaver. Only the case with ideal channel information (CSI) will be considered, that is $\rho_k^i$ is known for every $k$ and $i$.

A bit metric is calculated before carrying out the decoding. Let $l^i(x)$ denote the $i$-th label bit of $x$. From [3], [4], the optimum bit metric with ideal CSI for $l^i(x) = b$ at the time $k$ is given by

$$\tilde{\lambda}_k^i = \log \sum_{z \in \Upsilon_b^i} p_{\theta_k}(y_k|z) - \log \sum_{z \in \Upsilon_{\bar{b}}^i} p_{\theta_k}(y_k|z), \qquad (1)$$

Manuscript received August 17, 2004; accepted June 13, 2006. The associate editor coordinating the review of this letter and approving it for publication was K. Narayanan.

The authors are with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, 300, Taiwan, ROC (e-mail: benson_lw@realtek.com.tw; whsheen@mail.nctu.edu.tw; toshiba.cm93g@nctu.edu.tw).

Digital Object Identifier 10.1109/TWC.2006.04561

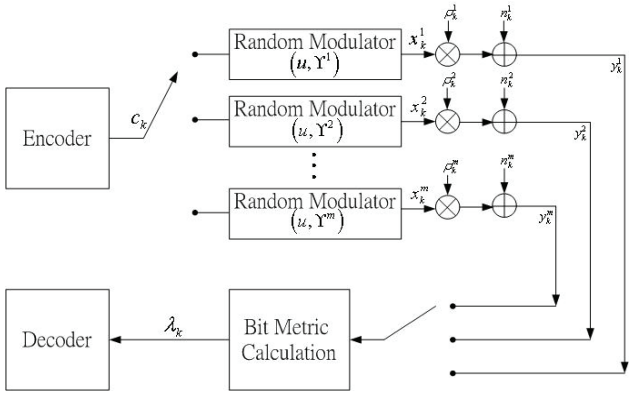


Fig. 1. Equivalent parallel independent channel model for BICM.

where $p(y|x)$ is the probability density function of $y$ given that $x$ is transmitted, and $\bar{b}$ is the complement of $b$. In practice, the branch metrics (1) is often simplified [3], [4] as

$$\lambda_k^i = \max_{z \in \Upsilon_b^i} \{\log p_{\theta_k}(y_k|z)\} - \max_{z \in \Upsilon_{\bar{b}}^i} \{\log p_{\theta_k}(y_k|z)\}, \tag{2}$$

which provides a good approximation to the optimal one at high SNRs. After calculating the bit metric, the Viterbi algorithm can then be used for decoding. The simplified bit metric (2) will be employed in this study.

## III. BER Upper Bounds

The BICM union and expurgated bounds in [4] will be reviewed first and then a new upper bound will be derived.

### A. BICM Union ($BICM_{UB}$) and Expurgated Bounds ($BICM_{EX}$) [4])

Consider a $(\kappa, n)$ convolutonal code, an upper bound on BER is given by [3], [4]

$$p_b \le \frac{1}{\kappa} \sum_{d=d_{min}}^{\infty} W_I(d) f(d, u, \Upsilon) \tag{3}$$

where $W_I(d)$ is the total weight of error events at Hamming distance $d$, $d_{min}$ the minimum Hamming distance of the code, and $f(d, u, \Upsilon)$ the pair-wise error probability (PEP). Let $d$ be the Hamming distance between the correct code sequence $\underline{c}$ and the error code sequence $\underline{\hat{c}}$. From [4], the PEP is given by

$$f(d, u, \Upsilon) = E_{\underline{S},\underline{U}} \left[ P(\underline{c} \to \underline{\hat{c}} | \underline{S}, \underline{U}) \right] \tag{4}$$

where $\underline{S} = (\cdots, S_{-1}, S_0, S_1, \cdots)$ and $\underline{U} = (\cdots, U_{-1}, U_0, U_1, \cdots)$ are the sequences of random variables to denote the operations of random label-position mapping and *symmetrization* of the use of the signal constellation, respectively, and $E[\cdot]$ denotes the operation of taking expectation. $S_k = i$, $i = 1, \cdots, m$ denotes that the coded bit $c_k$ is mapped to the $i$-th label position, $U_k = 1$ denotes that the mapping $u$ is used for $c_k$, and $U_k = 0$, the complement mapping $\bar{u}$ is used instead. $S_k$ and $U_k$ are assumed uniformly distributed, independent of each other and independent of other random variables. The parameter $\underline{U}$ was introduced in [4] for the purpose of easy analysis.

Let

$$\begin{aligned} \Upsilon_{\underline{c}}^{\underline{S}} &= \Upsilon_{c_1}^{i_1} \times \cdots \times \Upsilon_{c_k}^{i_k} \times \cdots \times \Upsilon_{c_d}^{i_d} \\ \Upsilon_{\underline{\hat{c}}}^{\underline{S}} &= \Upsilon_{\hat{c}_1}^{i_1} \times \cdots \times \Upsilon_{\hat{c}_k}^{i_k} \times \cdots \times \Upsilon_{\hat{c}_d}^{i_d} \end{aligned} \tag{5}$$

be the Cartesian product of the signal subsets $\Upsilon_{c_k}^{i_k}$ and $\Upsilon_{\hat{c}_k}^{i_k}$, selected by the bits $c_k$ and $\hat{c}_k$ with the label position $S_k = i_k$, respectively. (There is no loss of generality by just considering the consecutive errors in (5).) According to the bit metric $\lambda_k^i$ in (2), the path metric difference between $\underline{c}$ and $\underline{\hat{c}}$ is [4]

$$\begin{aligned} \delta &= \sum_{k=1}^{d} \max_{z_k \in \Upsilon_{c_k}^{i_k}} \log p(y_k|z_k) - \sum_{k=1}^{d} \max_{z_k \in \Upsilon_{\hat{c}_k}^{i_k}} \log p(y_k|z_k) \\ &= \max_{z_k \in \Upsilon_{c_k}^{i_k}} \log p(y_k|z_k) - \max_{z_k \in \Upsilon_{\hat{c}_k}^{i_k}} \log p(y_k|z_k) \end{aligned} \tag{6}$$

That results in

$$\begin{aligned} &P(\underline{c} \to \underline{\hat{c}} | \underline{S}, \underline{U}) = P(\delta < 0 | \underline{S}, \underline{U}) \\ &= E_{\underline{x}} \left[ P \left( \max_{\underline{z} \in \Upsilon_{\underline{c}}^{\underline{S}}} p(\underline{y}|\underline{z}) < \max_{\underline{z} \in \Upsilon_{\underline{\hat{c}}}^{\underline{S}}} p(\underline{y}|\underline{z}) | \underline{S}, \underline{U}, \underline{x} \right) \right] \\ &\le E_{\underline{x}} \left[ P \left( p(\underline{y}|\underline{x}) \le \max_{\underline{z} \in \Upsilon_{\underline{\hat{c}}_k}^{\underline{S}}} p(\underline{y}|\underline{z}) | \underline{S}, \underline{U}, \underline{x} \right) \right] \qquad (7) \\ &\le E_{\underline{x}} \left[ \sum_{\underline{z} \in \Upsilon_{\underline{\hat{c}}}^{\underline{S}}} P \left( p(\underline{y}|\underline{x}) \le p(\underline{y}|\underline{z}) | \underline{S}, \underline{U}, \underline{x} \right) \right] \qquad (8) \end{aligned}$$

where $\underline{x}$ is the transmitted signal sequence and $\underline{y}$ is the received one. The upper bound (7) is obtained because

$$p(\underline{y}|\underline{x}) \le \max_{\underline{z} \in \Upsilon_{\underline{\hat{c}}}^{\underline{S}}} p(\underline{y}|\underline{z})$$

and the upper bound (8) is due to the application of union bound. By using (7) and (8), the $BICM_{UB}$ in [4] is obtained as in (3) with $f(d, u, \Upsilon)$ replaced by

$$\begin{aligned} &f_{ub}(d, u, \Upsilon) \\ &= m^{-d} \sum_{\underline{S}} 2^{-d} \sum_{\underline{U}} 2^{-d(m-1)} \sum_{\underline{x} \in \Upsilon_{\underline{c}}^{\underline{S}}} \sum_{\underline{z} \in \Upsilon_{\underline{\hat{c}}}^{\underline{S}}} P(\underline{x} \to \underline{z}) \end{aligned} \tag{9}$$

Several observations can be made on $\text{BICM}_{\text{UB}}$. Firstly, the bound is quite general and can be applied to any mapping $u$ and constellation $\Upsilon$. Secondly, because of a union bound is invoked in (8), it is quite loose in general. And, finally, (9) is difficult to evaluate for a large $d$. In [4], by introducing the *symmetrization* operation , a much simpler method based on Laplace transform was proposed to ease this computation complexity. Furthermore, an approximation, called BICM expurgated bound $\text{BICM}_{\text{EX}}$, was also proposed in [4] by just including the unique nearest neighbor of $\underline{\hat{z}}$ of $\underline{x}$ in (9). With this simplification, $\text{BICM}_{\text{EX}}^{\text{M}}$ is obtained with $f(d, u, \upsilon)$ replaced by

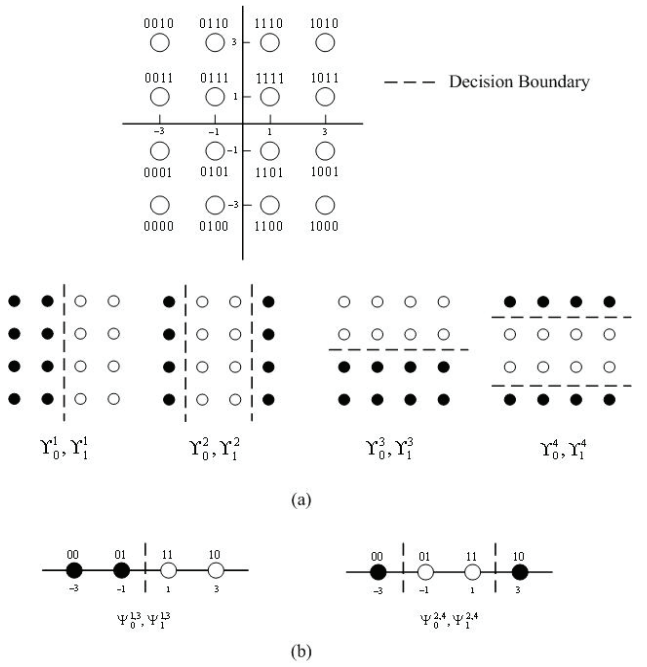


Fig. 2. (a) 16-QAM constellation with gray labeling, (b) equivalent one-dimensional constellation.

$$f_{ex}(d,u,\Upsilon) = m^{-d}\sum_{\underline{S}} 2^{-d}\sum_{\underline{U}} 2^{-d(m-1)}\sum_{\underline{x}\in\Upsilon_{\underline{c}}^{\underline{S}}} P(\underline{x}\to\underline{\hat{z}}). \quad (10)$$

$\text{BICM}_{\text{EX}}$ is claimed to be an upper bound in [4] for the square QAM constellation with gray labeling and an approximation for other forms of mappings and/or constellations. As is to be shown, however, theoretically it is an upper bound only for the QPSK case.

### B. New Upper Bound

In this section, a new easy-to-calculate, tight upper bound is derived for the square QAM constellation with gray labeling, which constitutes a very large portion of practical applications of the BICM systems. In the following, the 16 QAM case will be used to illustrate the derivation. Applications of this method to other sizes of constellation can be found in [9]. Prior to the derivation, we have the following observations.

*Observation 1*: With the bit metric given in (2), the two-dimensional square QAM constellation can be decomposed into two one-dimensional PAM ones and that facilitates the performance analysis. Fig. 2(b) shows the 16QAM example, where $\Psi_{c_k}^{i_k}$ is to denote the equivalent one-dimensional signal set of $\Upsilon_{c_k}^{i_k}$.

*Observation 2*: For $x \in \Psi_{c_k}^{i_k}$ and $\hat{x} \in \Psi_{\hat{c}_k}^{i_k}$, the pair-wise error probability $P\{x \to \hat{x}\}$ may not be evaluated as $P\{p(y|x) \le p(y|\hat{x})\}$ since the decision boundary is not necessarily equal to $\frac{(x+\hat{x})}{2}$. See $x=-1$, and $\hat{x}=3$ in Fig. 2(b) for an example. In this case, taking $P\{p(y|x) \le p(y|\hat{x})\}$ as the pair-wise error probability is amount to shift the decision boundary from $x=2$ to $x=1$, and that results in a looser upper bound. To circumvent this problem, we define $w(x,\hat{x})$ as the virtual signal associated with the pair $(x,\hat{x})$ such that $P\{x\to\hat{x}\} = P\{p(y|x) \le p(y|w(x,\hat{x}))\}$. For example, for $\Psi^{2,4}$ in Fig. 2(b), if $x=-1$ and $x=3$, then $w(x,\hat{x})=5$. With this arrangement, we can have a tighter bound because the true pair-wise error probability is always evaluated, as to be shown in (10). Note that it is not necessarily that $w(x,\hat{x})$ coincides with one of the constellation points. For notation simplicity, $w(x,\hat{x})$ will be denoted as in the sequel.

Equipped with these observations, we can now proceed to derive the new upper bound. Following the notation in (6), we have

$$\begin{aligned} P(\underline{c}\to\underline{\hat{c}}|\underline{S},\underline{U}) &= P(\delta<0|\underline{S},\underline{U}) \\ &= E_{\underline{x}}\left[P\left(\max_{\underline{z}\in\Upsilon_{\underline{c}}^{\underline{S}}} p(\underline{y}|\underline{z}) < \max_{\underline{z}\in\Upsilon_{\underline{\hat{c}}}^{\underline{S}}} p(\underline{y}|\underline{z})|\underline{S},\underline{U},\underline{x}\right)\right] \\ &\le E_{\underline{x}}\left[P\left(p(\underline{y}|\underline{x}) \le \max_{\underline{w}\in\Omega_{\underline{\hat{c}}_k}^{\underline{S}}} p(\underline{y}|\underline{z})|\underline{S},\underline{U},\underline{x}\right)\right] \quad (11) \\ &\le E_{\underline{x}}\left[\sum_{\underline{w}\in\Omega_{\underline{\hat{c}}}^{\underline{S}}} P\left(p(\underline{y}|\underline{x}) \le p(\underline{y}|\underline{z})|\underline{S},\underline{U},\underline{x}\right)\right] \quad (12) \end{aligned}$$

where

$$\Omega_{\underline{\hat{c}}}^{\underline{S}} = \{\underline{w} = (\cdots, w_{-1}, w_0, w_1, \cdots)|w_k = \text{the virtual signal associated with}(x_k,\hat{x}_k).\}. \quad (13)$$

The bound (12) is tighter than (8) because of the true decision boundary is used in (11). For example, see Fig. 2(b) with $\Psi^{2,4}$ , if $x=-1$, and $\hat{x}=3$, then by using (8), the decision boundary is $x=1$, but in fact the true boundary is $x=2$.

The upper bound in (12) can be made tighter by removing the irrelevant error events in the union bound. Since we just need to consider the one-dimensional constellation, for a given $x$, only the nearest signals on both side of $x$, denoted $\hat{x}_l$ and $\hat{x}_r$, in $\Psi_{\hat{c}_k}^{S_k}$ required to be included, Of course if error can occur only one-sided, there is no need to consider both sides. It is easy to see that the set of $y$ that results in an error signal $\hat{x}\neq\hat{x}_l$ and $\hat{x}\neq\hat{x}_r$ have been accounted for when considering $\hat{x}_l$ or $\hat{x}_r$. Let $\hat{\Omega}_{\underline{\hat{c}}}^{\underline{S}}$ be the subset of $\Omega_{\underline{\hat{c}}}^{\underline{S}}$ with the virtual signals of $\hat{x}_l$ and $\hat{x}_r$ being retained. Then a tighter bound is obtained for the conditional PEP

$$P(\underline{c}\to\underline{\hat{c}}|\underline{S},\underline{U}) \le E_{\underline{x}}\left[\sum_{\underline{w}\in\hat{\Omega}_{\underline{\hat{c}}}^{\underline{S}}} P\left(p(\underline{y}|\underline{x}) \le p(\underline{y}|\underline{w})|\underline{S},\underline{U},\underline{x}\right)\right] \quad (14)$$

The idea of expurgating the irrelevant error events in (14) is similar to the one used for $\text{BICM}_{\text{EX}}$ in [4]. Nevertheless, only one side of the error signals or is included in $\text{BICM}_{\text{EX}}$ and that makes it an upper bound only for the case of QPSK. By using (14), the new upper bound $\text{BICM}_{\text{NEW}}$ is obtained in (3) with $f(d,u,\Upsilon)$ replaced by

$$f_{new}(d,u,\Upsilon) = m^{-d}2^{-md}\sum_{\underline{S}}\sum_{\underline{U}}\sum_{\underline{x}\in\Psi_{\underline{c}}^{\underline{S}}}\sum_{\underline{w}\in\hat{\Omega}_{\underline{x}}^{\underline{S}}} P(\underline{x}\to\underline{\hat{w}}). \quad (15)$$

Methods for computing the weight distribution $W_I(d)$ for punctured convolutional codes are well known as in [7]. Only the calculation of $f_{new}(d,u,\Upsilon)$ will be discussed in the following.

### C. Pair-Wise Error Probability

For AWGN channels, $P(\underline{x}\rightarrow\underline{w})$ is given by

$$P(\underline{x}\rightarrow\underline{w})=Q\left(\rho(\underline{x};\underline{w})\sqrt{\frac{2m\Delta^2RE_b}{dN_0}}\right) \tag{16}$$

where

$$Q(\alpha)=\frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\infty}e^{-(\frac{y^2}{2})}dy, \text{and}$$

$$\rho(\underline{x};\underline{w})=\sum_{1}^{d}\frac{|x_k-w_k|}{2}$$

is a half of the Euclidean distance between $\underline{w}$ and $\underline{x}$, $\Delta$ the power normalization factor of QAM constellation, $E_b$ bit energy, $N_0$ the one-sided noise density and $R$ code rate. $\Delta=\frac{1}{\sqrt{10}}$for 16QAM. Using (16), (15) becomes

$$\begin{aligned}&f_{new}(d,u,\Upsilon)\\&=m^{-d}2^{-md}\sum_{\underline{S}}\sum_{\underline{U}}\sum_{\underline{x}\in\Psi_{\underline{c}}^{\underline{s}}}\\&\quad\sum_{\underline{w}\in\hat{\Omega}_{\underline{x}}^{\underline{S}}}Q\left(\rho(\underline{x};\underline{w})\sqrt{\frac{2m\Delta^2RE_b}{dN_0}}\right).\end{aligned} \tag{17}$$

The computation of (17) can be further simplified by utilizing symmetrization constellation of one-dimensional PAM constellation. Consider 16QAM as an example. Define the signal subsets $X_1\triangleq\{(0,0),(1,0)\}$ and $X_2\triangleq\{(0,1),(1,1)\}$. From Fig. 2(b) it is observed that for every $x\in X_1$, the pair-wise error probability $P(x\rightarrow w)$ is the same, and so does for $x\in X_2$, no matter if the MSB (first label position) or the LSB (second label position) is considered. Indeed, this is the case; for MSB, $\rho(x;w)=1$, $x\in X_2$ , and $\rho(x;w)=3$, $x\in X_1$. Likewise, for LSB, $\rho(x;w)=1$, $x\in X_1$ and $\rho(x;w_1)=3$, $\rho(x;w_2)=1$ for $x\in X_2$. Note that for $x\in X_2$, there are two regions for errors to occur when the bit error in LSB is considered. With this symmetry property, (17) can be simplified as follows.

$$\begin{aligned}&f_{new}(d,u,\Upsilon)\\&=2^{-2d}\sum_{i=0}^{d}\binom{d}{i}\sum_{i_1=0}^{i}\binom{i}{i_1}\sum_{j_1=0}^{j}\binom{j}{j_1}\sum_{s=0}^{j_2}\binom{j_2}{s}\\&\quad Q\left(\{3i_2+i_2+j_1+[(j_2-s)+3s]\}\sqrt{\frac{4RE_b}{5dN_0}}\right)\end{aligned} \tag{18}$$

where $d=i+j$ is the total number of bit errors, $i$ and $j$ are the number of bit errors in MSB and LSB, respectively, $i=i_1+i_2$, $i_1$ is the number of MSB errors with $x\in X_1$, $i_2$ is with $x\in X_2$, $j=j_1+j_2$, $j_1$ is the number of LSB errors with $x\in X_1$, $j_2$ is with $x\in X_2$, $s$ is the number of LSB errors with $x\in X_2$ and $\rho(x;w_1)=3$. The Laplace transform method in [4] can be equally applied to the evaluation of (17).

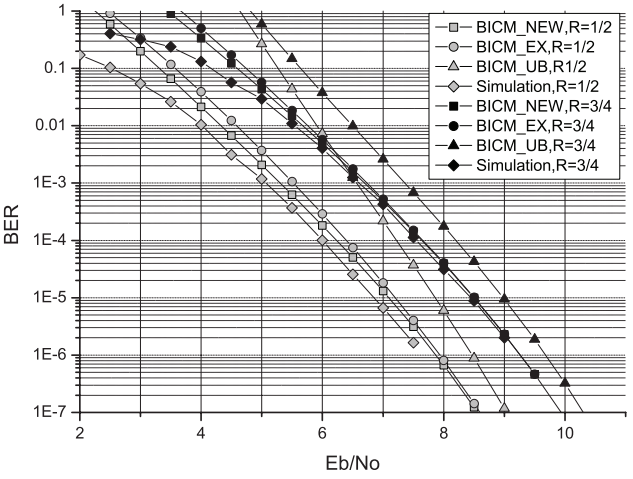


Fig. 3. Comparisons of different bounds for 16 QAM in AWGN channels.

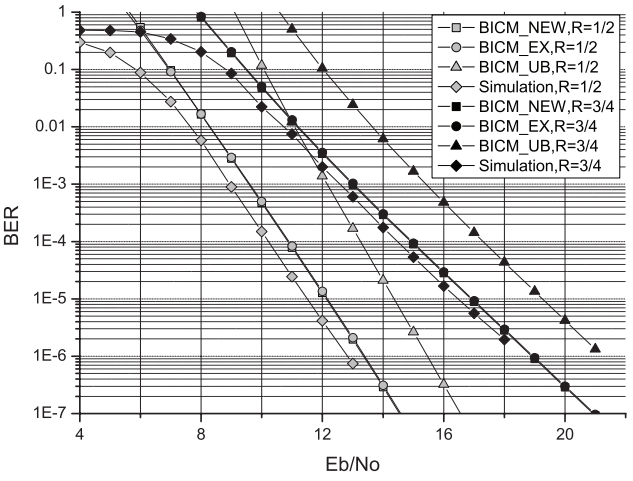


Fig. 4. Comparisons of different bounds for 16QAM in Rayleigh fading channels.

For fading channels, we apply the Chernoff bound [8] to obtain

$$P(\underline{x}\rightarrow\underline{w})\leq\frac{1}{2}\prod_{i=1}^{d}\frac{1}{1+SNR(\frac{x_i-w_i}{2})^2}, \tag{19}$$

Therefore, PEP is upper-bounded by

$$\begin{aligned}&f_{new}(d,u,\Upsilon)\\&=m^{-d}2^{-md}\sum_{\underline{S}}\sum_{\underline{U}}\\&\quad\sum_{\underline{x}\in\Psi_{\underline{c}}^{\underline{S}}}\sum_{\underline{w}\in\hat{\Omega}_{\underline{x}}^{\underline{S}}}\frac{1}{2}\prod_{i=1}^{d}\frac{1}{1+SNR(\frac{x_i-w_i}{2})^2},\end{aligned} \tag{20}$$

where $SNR=\frac{m\Delta^2RE_b}{N_0}$. Again, by using the symmetry property, we have for 16-QAM

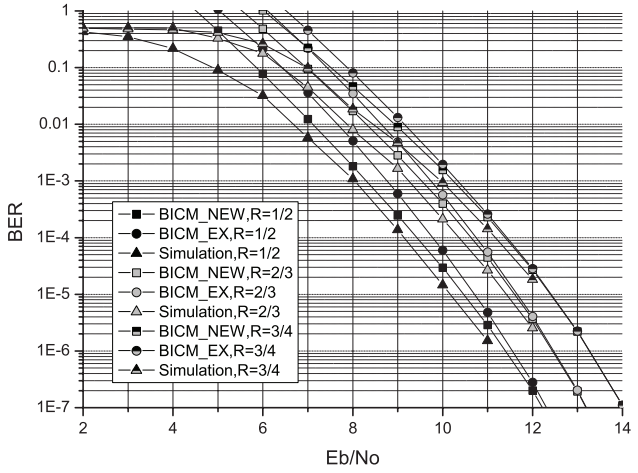


Fig. 5. Performance of BICM in AWGN channels (64QAM).

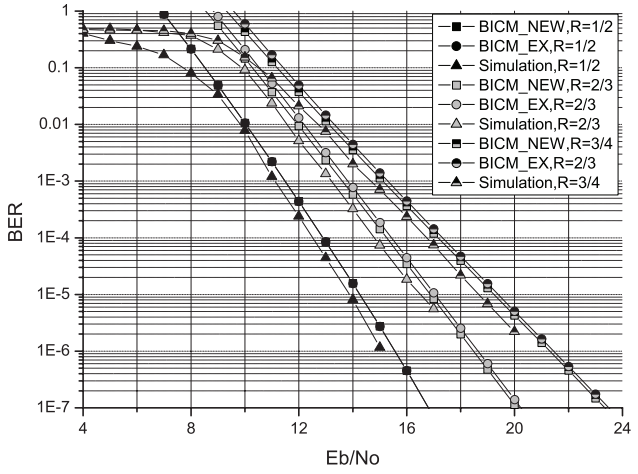


Fig. 6. Performance of BICM in Rayleigh fading channels (64QAM).

$$f_{new}(d,u,\Upsilon) = 2^{-2d}\sum_{i=0}^{d}\binom{d}{i}\sum_{i_1=0}^{i}\binom{i}{i_1}\sum_{j_1=0}^{j}\binom{j}{j_1}\sum_{s=0}^{j_2}\binom{j_2}{s} \frac{1}{2}\left(\frac{1}{1+3^2SNR}\right)^{i_i+s}\left(\frac{1}{1+SNR}\right)^{i_2+j_1+j_2-s} \quad (21)$$

where $SNR=\frac{2RE_b}{5N_0}$

## IV. Numerical Results

The BICM based on the $\frac{1}{2}$-rate convolutional code, with the generator polynomials (133,171), is used for numerical examples. The punctured patterns of different code rates are adopted from [7]. More numerical results can be found in [9]. Figs. 3 and 4 compare $\text{BICM}_{\text{UB}}$, $\text{BICM}_{\text{EX}}$ and $\text{BICM}_{\text{NEW}}$ for 16QAM over AWGN and fading channels, respectively. As is evident, $\text{BICM}_{\text{UB}}$ is quite a loose bound. $\text{BICM}_{\text{EX}}$ is a good approximation, although theoretically it is not an upper bound. As is expected, the new $\text{BICM}_{\text{NEW}}$ is very tight for BERs of practical interest. Figs. 5 and 6 give the numerical results with 64QAM. Again, $\text{BICM}_{\text{NEW}}$ provides a very tight upper bound for all the cases.

## V. Conclusion

This paper investigates the BER performance of BICM systems. A new BER upper bound is derived for BICM with the square QAM and gray labeling. Numerical results show that the proposed bound is tighter than the well-known BICM union in [4].

## References

[1] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, pp. 56–67, Jan. 1982.

[2] D. Divsalar and M. Simon, "The design of trellis coded MPSK for fading channel: Performance criteria," *IEEE Trans. Commun.*, vol. 36, pp. 1004–1012, Sep. 1988.

[3] E. Zehavi, "Eight-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.

[4] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–945, May 1998.

[5] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 944–957, May 2001.

[6] R. Visoz and A. O. Berthet, "Iterative decoding and channel estimation for space-time BICM over MIMO block fading multipath AWGN channel," *IEEE Trans. Commun.*, vol. 51, pp. 1358–1367, Aug. 2003.

[7] D. Haccoun and G. Begin, "High-rate punctured convolutional codes for Viterbi and sequential decoding," *IEEE Trans. Commun.*, vol. 37, pp. 1113–1125, Nov. 1989.

[8] C. Tellambura, "Evaluation of the exact union bound for trellis-coded modulations over fading channels," *IEEE Trans. Commun.*, vol. 44, pp. 1693–1699, Dec. 1996.

[9] Liang-Wei Huang, "Tight Performance Upper Bounds for Bit Interleaved Coded QAM Modulation with Gray Labeling." MS thesis, National Chiao Tung University, Hsinchu, Taiwan, Aug. 2004.