

On Multilevel Coded MPSK for the Rayleigh Fading Channel

Ruey-Yi Wei and Mao-Chao Lin

Department of Electrical Engineering, National Taiwan University, ROC

Abstract: We derive a new upperbound on pairwise error probability for MPSK sequences over the Rayleigh fading channel when CSI is not available. This is obtained by adding weight factors in calculating the symbol metrics. Simulation shows that the weight factors which optimize the upperbound are likely to optimize the error rate too. We also design multilevel coded MPSK schemes which are suitable for the Rayleigh fading channel. These schemes are basically in the form of block coded modulation with interblock memory (BCMIM) [4], which was originally designed for additive white Gaussian noise (AWGN) channel. The added weight factors are found to be effective in improving the error performances of these BCMIM schemes in case that CSI is not available.

I Introduction

In 1982, Ungerboeck[1] introduced the concept of coded modulation which was originally designed for the additive white Gaussian noise (AWGN) channel. For a coded modulation system applied to the AWGN channel, it is desired to maximize its minimum squared Euclidean distance (MSED). In [2], Divsalar and Simon investigate the situation of applying a coded MPSK system to the Rayleigh fading channel. Through the derivation of upperbounds on the pairwise error probability of MPSK sequences, it has been shown that maximizing both minimum symbol distance (MSD) and minimum product distance (MPD) is desired regardless of whether channel state information (CSI) is available or not. In this paper, we derive a better upperbound on the pairwise error probability for the Rayleigh fading channel without CSI by adding weight factors in calculating the symbol metrics. The channel under consideration in this paper is the frequency nonselective slow Rayleigh fading channel with perfect phase tracking. We find weight factors which can optimize the upperbound. The new upperbound also shows that maximizing MSD and MPD is desired. Simulation on several examples show that the weight factors which optimize the upperbound are likely to be the best choices in lowering the error rates.

Multilevel coding is a useful method to construct coded modulation systems [3]. The basic idea of multilevel coding is to partition a signal set into several levels and to separately encode each level with a proper component code. We can use either a block code or a convolutional code as the component code in each level. If the component codes

for all the levels are block codes, it is a kind of block coded modulation (BCM).

In [4], a kind of multilevel coded modulation called block coded modulation with interblock memory (BCMIM) was designed for the AWGN channel to increase the coding rate without decreasing MSED as compared to the associated multilevel BCM. In this paper, we modify the design of BCMIM for MPSK signals applied to the Rayleigh fading channel such that the coding rate can be increased without decreasing MSD and MPD. The error performances of these BCMIM are simulated by computer. When CSI is not available, the error performances of the designed BCMIM schemes are not as good as we expect if conventional symbol metrics are used. However, by adding proper weight factors to the symbol metrics, improvement of the error performances can be observed. The optimal choice for the weight factor is a value proportional to the inverse of the level Euclidean distance in each level. In this paper, the coding rate, error performances and decoding complexities are provided for some BCM and BCMIM examples so that a fair comparison can be made.

II Upperbounds on the Error Rate for Coded MPSK

Let $\mathbf{x} = (x_1, x_2, \dots, x_N)$ be a transmitted coded MPSK symbol sequence of length N , where x_i is the transmitted symbol at time i . Let $\mathbf{r} = (r_1, r_2, \dots, r_N)$ be the corresponding received sequence, where the i th element, r_i , is given by

$$r_i = \rho_i x_i + n_i, \quad (1)$$

where n_i is a zero mean complex Gaussian variable with variance σ_n^2 and ρ_i is a normalized random variable with Rayleigh distribution.

Consider the case that ideal CSI is available. Given the condition that the fading gain sequence is $\rho = (\rho_1, \dots, \rho_N)$, the pairwise error probability of choosing $\hat{\mathbf{x}}$ instead of \mathbf{x} is

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \rho\} = Pr\left\{\sum_{i=1}^N [m(r_i, x_i, \rho_i) - m(r_i, \hat{x}_i, \rho_i)] > 0\right\}, \quad (2)$$

where $m(r_i, x_i, \rho_i)$ is the metric of x_i which will be used in the decoding trellis.

By setting $m(r_i, x_i, \rho_i) = |r_i - \rho_i x_i|^2$, it [2] can be shown that for reasonably large $\frac{E_s}{N_0}$ values, the pairwise

error probability of choosing $\hat{\mathbf{x}}$ instead of \mathbf{x} is bounded by

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\} \leq \left(\prod_{i \in \eta} \frac{E_s}{4N_0} |x_i - \hat{x}_i|^2 \right)^{-1}. \quad (3)$$

This upperbound was derived by Divsalar and Simon [2].

Now we consider the case that CSI is absent, the pairwise error probability of choosing $\hat{\mathbf{x}}$ instead of \mathbf{x} given the condition of fading gain sequence $\boldsymbol{\rho}$ is

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \boldsymbol{\rho}\} = Pr\left\{ \sum_{i=1}^N [m(r_i, x_i) - m(r_i, \hat{x}_i)] > 0 \right\}. \quad (4)$$

Traditionally, $m(r_i, x_i)$ is set to be $|r_i - x_i|^2$. We propose to set $m(r_i, x_i) = w_i |r_i - x_i|^2$, where w_i is a weight factor. By applying Chernoff bound to (4), we have

$$\begin{aligned} Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \boldsymbol{\rho}\} &\leq \prod_{i \in \eta} E\{ \exp[\lambda(m(r_i, x_i) - m(r_i, \hat{x}_i))] \}. \\ &= \prod_{i \in \eta} \exp[-\lambda w_i |x_i - \hat{x}_i|^2 - 2\lambda w_i (\rho_i - 1) \\ &\quad \times Re\{x_i(x_i - \hat{x}_i)^*\}] \\ &\quad \times E\{ \exp[-2\lambda w_i Re\{n_i(x_i - \hat{x}_i)^*\}] \}, \quad (5) \end{aligned}$$

where y^* is the complex conjugate of y and $Re\{y\}$ is the real part of y . Suppose that the $|x_i| = |\hat{x}_i|$, it can be shown that $|x_i - \hat{x}_i|^2 = 2Re\{x_i(x_i - \hat{x}_i)^*\}$. and $E\{ \exp[-2\lambda w_i Re\{n_i(x_i - \hat{x}_i)^*\}] \} = \exp[\lambda^2 w_i^2 \sigma_n^2 |x_i - \hat{x}_i|^2]$. With $d_i^2 = |x_i - \hat{x}_i|^2$, we have

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \boldsymbol{\rho}\} \leq \prod_{i \in \eta} \exp[-\lambda w_i \rho_i d_i^2 + \lambda^2 w_i^2 \sigma_n^2 d_i^2]. \quad (6)$$

Normalizing the Chernoff parameter (replacing λ by $\lambda \sigma_n^2$) and substituting $\frac{E_s}{N_0}$ for $\frac{1}{\sigma_n^2}$ in (6), we have

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \boldsymbol{\rho}\} \leq \prod_{i \in \eta} \exp[\lambda^2 w_i^2 \frac{E_s}{N_0} d_i^2] \exp[-\lambda w_i \rho_i \frac{E_s}{N_0} d_i^2]. \quad (7)$$

With the assumption of infinite interleaving, the pairwise error probability is

$$\begin{aligned} Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\} &= \int_0^\infty \cdots \int_0^\infty Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}} \mid \boldsymbol{\rho}\} p(\rho_1) p(\rho_2) \\ &\quad \cdots p(\rho_N) d\rho_1 d\rho_2 \cdots d\rho_N, \quad (8) \end{aligned}$$

where ρ_i and ρ_j are identical and mutually independent for $i \neq j$.

With a procedure similar to that used in [2], we can have

$$\begin{aligned} Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\} &\leq \prod_{i \in \eta} \exp[\lambda^2 w_i^2 \frac{E_s}{N_0} d_i^2] \times [1 - \\ &\quad \frac{1}{\sqrt{\pi}} \int_0^\pi \eta(\theta) \exp(\eta^2(\theta)) \operatorname{erfc} \eta(\theta) d\theta] \quad (9) \end{aligned}$$

where

$$\eta(\theta) = \lambda w_i \frac{E_s}{2N_0} d_i^2$$

Since $\operatorname{erfc} x \approx \frac{\exp(-x^2)}{\sqrt{\pi}x} (1 - \frac{1}{2x^2})$, we have

$$\begin{aligned} Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\} &\leq \prod_{i \in \eta} \frac{\exp[\lambda^2 w_i^2 \frac{E_s}{N_0} d_i^2]}{2\lambda^2 w_i^2 (\frac{E_s}{2N_0})^2 d_i^2} \\ &= \frac{\exp[\lambda^2 \frac{E_s}{N_0} \sum_{n \in \eta} w_n^2 d_n^2]}{2^{L_\eta} \lambda^{2L_\eta} (\frac{E_s}{2N_0})^{2L_\eta} \prod_{i \in \eta} w_i^2 d_i^2}, \quad (10) \end{aligned}$$

where L_η is the cardinality of η .

To minimize the righthand side of (10), we need

$$w_j = (\lambda \frac{E_s}{N_0} d_j)^{-1} \propto \frac{1}{d_j}, \quad (11)$$

for each $j \in \eta$. It follows from (10), (11) and

$$\lambda^2 = \frac{L_\eta}{\frac{E_s}{N_0} \sum_{n \in \eta} w_n^2 d_n^2},$$

we have

$$Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\} \leq \frac{(2e)^{L_\eta}}{(\frac{E_s}{N_0})^{L_\eta}} \left(\prod_{i \in \eta} d_i^2 \right)^{-1} \quad (12)$$

To see the effect of the weight factor, we give two simple examples as follows.

Example 1 : Consider two 8PSK sequences, $\mathbf{x} = (x_1, x_2, x_3)$ and $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)$, such that $d_1^2 = 0.586$, $d_2^2 = 2$, $d_3^2 = 4$. The ratio of optimal weight factors is $w_1 : w_2 : w_3 = \frac{1}{\sqrt{0.586}} : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{4}} = 1 : 0.5412 : 0.3827$. Upperbounds and simulation results are shown in Fig. 1.

Example 2 : Consider two 16PSK sequences, $\mathbf{x} = (x_1, x_2, x_3, x_4)$ and $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)$, such that $d_1^2 = 0.152$, $d_2^2 = 0.586$, $d_3^2 = 2$, $d_4^2 = 4$. Upperbounds and simulation results are shown in Fig. 2.

Although to minimize the upperbound of pairwise error probability, we have shown that the best weight factor w_j is proportional to the inverse of d_j , we are unable to theoretically prove that to minimize the pairwise error probability $Pr\{\mathbf{x} \rightarrow \hat{\mathbf{x}}\}$ also requires that w_j be proportional to the inverse of d_j . However, simulation results indicate that probably setting w_j to be proportional to the inverse of d_j is the best choice. Hence, we suggest that : "for decoding in the Rayleigh fading channel, the best weight factor at the j -th position may be proportional to the inverse of the Euclidean distance between the two associated symbols".

III Several Coded MPSK Systems for the Rayleigh Fading Channel

In this section, we will design several coded MPSK systems which are suitable for the Rayleigh fading channel. For the convenience of presentation, we restrict M to be 8.

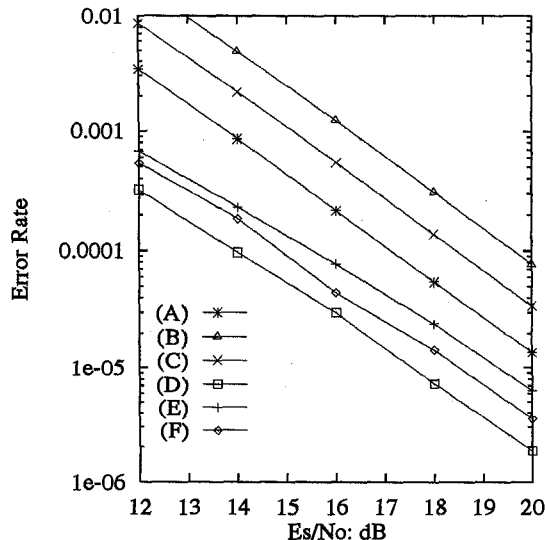


Figure 1: $Pr\{x \rightarrow \hat{x}\}$ for 8PSK sequences of Example 1.

- (A) Upperbounds with CSI.
- (B) Upperbounds without CSI using $w_1 : w_2 : w_3 = 1 : 1 : 1$.
- (C) Upperbounds without CSI using $w_1 : w_2 : w_3 = 1 : 0.5412 : 0.3827$.
- (D) Simulation results with CSI.
- (E) Simulation results without CSI using $w_1 : w_2 : w_3 = 1 : 1 : 1$.
- (F) Simulation results without CSI using $w_1 : w_2 : w_3 = 1 : 0.5412 : 0.3827$.

Consider an 8PSK signal constellation as given in [1], in which each signal point is labelled by (a, b, c) , where a, b , and $c \in \{0, 1\}$. By successive two-way partitions of 8PSK signal set, the intra-subset squared Euclidean distances are $\delta_a^2 = 0.586$, $\delta_b^2 = 2$ and $\delta_c^2 = 4$ respectively [1]. Let $(a_1, b_1, c_1), (a_2, b_2, c_2), \dots$ be a sequence of transmitted 8PSK signals. A conventional multilevel coding system is designed in such a way that (a_1, a_2, \dots) is a codeword of a binary code C_a , (b_1, b_2, \dots) is a codeword of a binary code C_b , and (c_1, c_2, \dots) is a codeword of a binary code C_c . Here, C_i is the code used for coding the level i , $i \in \{a, b, c\}$. Suppose that C_i is an (n, k_i, d_i) binary block code with generator matrix G_i , where $i \in \{a, b, c\}$. Then we have a BCM for which each block consists of n 8PSK signal points and the coding rate is $(k_1 + k_2 + k_3)/n$ bits per 8PSK signal point. Such a BCM can be easily decoded by a 3-stage decoding, where at the i -th stage of decoding, a trellis for C_i is used.

The important parameters which will affect the error performances of a coded MPSK for the Rayleigh fading channel are MSD, MPD and $N(\alpha, \beta)$, where $N(\alpha, \beta)$ is the number of neighbors at a symbol distance of α and at a product distance of β . From the upperbounds given in (3) and (12), we see that MSD is the most important parameter for high signal-to-noise ratio (SNR). Hence, it

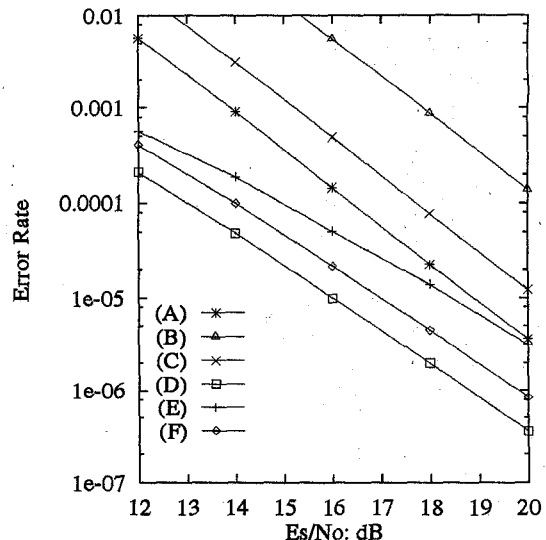


Figure 2: $Pr\{x \rightarrow \hat{x}\}$ for 16PSK sequences of Example 2.

- (A) Upperbounds with CSI.
- (B) Upperbounds without CSI using $w_1 : w_2 : w_3 : w_4 = 1 : 1 : 1 : 1$.
- (C) Upperbounds without CSI using $w_1 : w_2 : w_3 : w_4 = 1 : 0.5098 : 0.2759 : 0.1951$.
- (D) Simulation results with CSI.
- (E) Simulation results without CSI using $w_1 : w_2 : w_3 : w_4 = 1 : 1 : 1 : 1$.
- (F) Simulation results without CSI using $w_1 : w_2 : w_3 : w_4 = 1 : 0.5098 : 0.2759 : 0.1951$.

is natural to set C_a, C_b , and C_c to be the same code, to achieve a large MSD.

Example 3 (BCM-1): Let $C_a = C_b = C_c$ be an $(8, 4, 4)$ binary code. Then, we have a BCM. The coding rate is $12/8$ bits per symbol. The parameters MSD, MPD and $N(\text{MSD}, \text{MPD})$ are 4, $(0.586)^4$ and 224 respectively. Simulation results for the Rayleigh fading channel with CSI and without CSI which are obtained from a 3-stage decoding are respectively given in Fig. 4, Fig. 5 and Fig. 6, where in Fig. 4 weight factor is not used and in Fig. 5 optimal weight factors are used. The ratio of optimal weight factors for level a, b and c is $w_a : w_b : w_c = 1 : 0.5412 : 0.3827$. In each stage of the 3-stage decoding, a 4-state trellis for the $(8, 4, 4)$ binary code is needed.

For BCM-1, the bit error rate (BER) is not satisfactory for high SNR, since the MPD δ_a^2 is not large enough. One way to increase δ_a^2 is to decrease the code rate of C_a as given in Example 4.

Example 4 (BCM-2): Replace C_a in BCM-1 by an $(8, 1, 8)$ code. Then the coding rate is reduced to $9/8$ bits per symbol. The parameters MSD, MPD and $N(\text{MSD}, \text{MPD})$ are 4, 16 and 224 respectively. Simulation results given in Figures 4, 5 and 6 show the improvement of BER for high SNR. For both BCM-1 and BCM-2, the decoding results without CSI are independent of the weight factors Since

the coding of each level is independent of each levels.

The BCMIM scheme [4] proposed in 1994 is for the AWGN channel. For the Rayleigh fading channel, the design criterion is much different.

Let $(a_1, \dots, a_n, b_1, \dots, b_n, c_1, \dots, c_n)$ represent a block of n 8PSK signal points and $(a_1^+, \dots, a_n^+, b_1^+, \dots, b_n^+, c_1^+, \dots, c_n^+)$ represent the following block. These two adjacent blocks are combined to be called a superblock. For a BCM, we have $(a_1, \dots, a_n) = \tilde{u}_a \cdot G_a$, $(b_1, \dots, b_n) = \tilde{u}_b \cdot G_b$ and $(c_1, \dots, c_n) = \tilde{u}_c \cdot G_c$, where G_i is the generator matrix of the (n, k_i, d_i) binary block code C_i and \tilde{u}_i is a k_i -bit message for $i \in \{a, b, c\}$. For a BCMIM with interblock memory between levels a and b , it is designed such that

$$(b_1, \dots, b_n, a_1^+, \dots, a_n^+) = (\tilde{u}_r, \tilde{u}_a^+, \tilde{u}_b) \cdot G \quad (13)$$

where \tilde{u}_r is a k_r -bit message,

$$G = \begin{pmatrix} G_{br} & G_{ar} \\ 0 & G_a \\ G_b & 0 \end{pmatrix} \quad (14)$$

is the generator matrix of a $(2n, k_a + k_b + r)$ binary code C , G_{ar} and G_{br} are both $r \times n$ matrices. Also, coding for level c is obtained by $(c_1, \dots, c_n) = \tilde{u}_c \cdot G_c$. The coding configuration is given in Fig. 3.

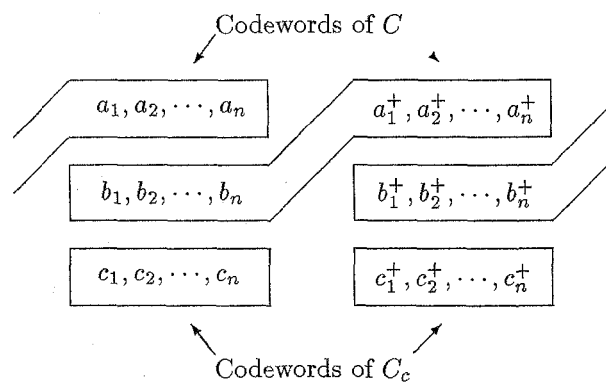


Figure 3: The structure of BCMIM with interblock memory between levels a and b .

In this way, the coding rate will be $(r + k_a + k_b + k_c)/n$ bits per 8PSK signal point. If we switch the roles of the "b" and "c" symbols in the above coding design, then we have a BCMIM with interblock memory between levels a and c . Of course, we can provide interblock memory between levels b and c in a similar way.

Consider the case that the interblock memory is provided between levels a and b . Let the MSD and MPD of the original BCM be δ_H and Δ_P^2 respectively. Let C_{ii} denote the $(n, k_i + r, d_{ii})$ binary block code with generator matrix $[G_{ir}^T, G_i^T]^T$, $i \in \{a, b, c\}$, where M^T is the transpose of a matrix M . Let $(\tilde{u}_r, \tilde{u}_a^+, \tilde{u}_b, \tilde{u}_c)$ and $(\tilde{u}'_r, \tilde{u}'_a^+, \tilde{u}'_b, \tilde{u}'_c)$ be distinct messages

which are encoded into $(b_1, \dots, b_n, c_1, \dots, c_n, a_1^+, \dots, a_n^+)$ and $(b'_1, \dots, b'_n, c'_1, \dots, c'_n, a_1'^+, \dots, a_n'^+)$ respectively. Consider the following conditions.

(i) Suppose that $\tilde{u}_r = \tilde{u}'_r$. Then, the MSD and MPD between the associated superblocks are equal to δ_H and Δ_P^2 respectively.

(ii) Suppose that $\tilde{u}_r \neq \tilde{u}'_r$. Then, the MSD and MPD between the associated superblocks are

$$\delta_{Hr} = d_{aa} + d_{bb} \quad (15)$$

and

$$\Delta_{Pr}^2 = (\delta_a^2)^{d_{aa}} (\delta_b^2)^{d_{bb}} \quad (16)$$

respectively. In our design, we require either (A) $\delta_{Hr} = \delta_H$ and $\Delta_{Pr}^2 > \Delta_P^2$ or (B) $\delta_{Hr} > \delta_H$. Note that for high SNR, symbol distance is more important than product distance. Hence, BCMIM and BCM will have almost the same bit error rate (BER) at the same E_s/N_o and BCMIM will have lower BER than BCM at the same E_b/N_o which results from the smaller bandwidth required by BCMIM, where E_s is the average energy of each signal point, E_b is the average energy of each message bit and N_o is the one-sided power spectral density of AWGN.

BCMIM for which interblock memory is provided between other levels can be similarly designed.

We will show specific designs of BCMIM for the Rayleigh fading channel in the following examples.

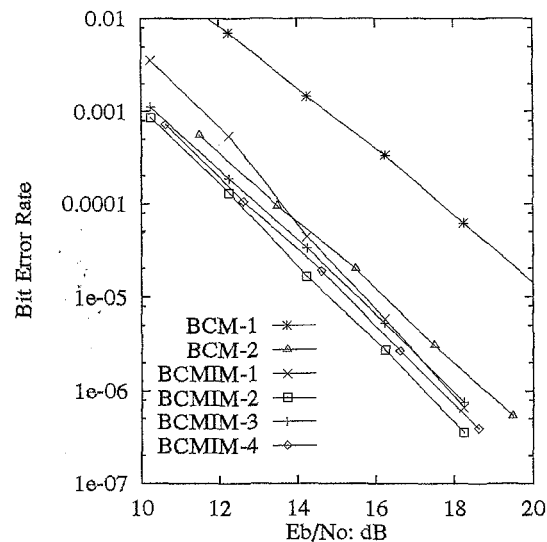


Figure 4: Simulations with CSI.

Example 5 (BCMIM-1) : For this BCMIM, C_a, C_b, C_c, C_{aa} and C_{bb} are $(8,1,8), (8,4,4), (8,4,4), (8,4,4)$ and $(8,7,2)$ binary codes respectively. The code rate is 12/8 bits per signal point. In the 2-stage decoding, an 8-state trellis for the $(16,8,4)$ binary code C and a 4-state trellis for C_c are needed.

Example 6 (BCMIM-2) : For this BCMIM, C_a, C_b, C_c, C_{aa} and C_{cc} are $(8,1,8), (8,4,4), (8,4,4), (8,4,4)$ and $(8,7,2)$

binary codes respectively. The code rate is 12/8 bits per signal point. In the 2-stage decoding, an 8-state trellis for the (16,8,4) binary code and a 4-state trellis for C_b are needed.

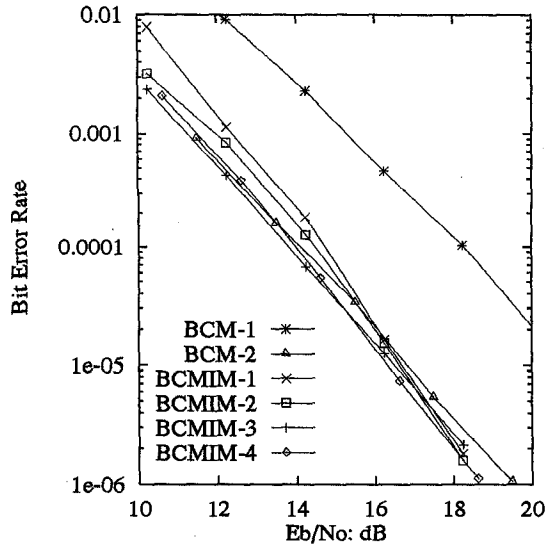


Figure 5: Simulations without CSI using $w_a : w_b : w_c = 1 : 1 : 1$.

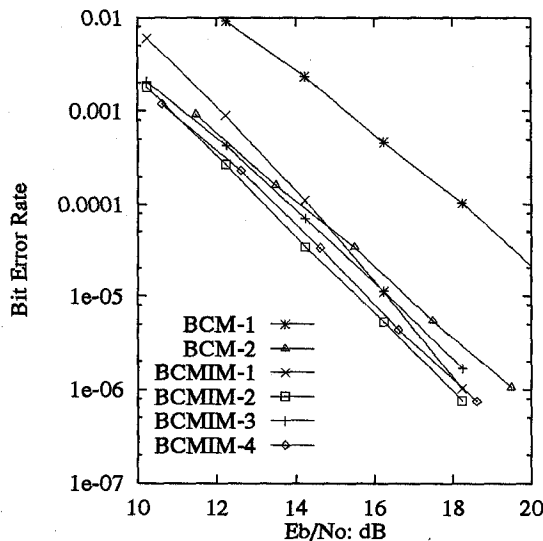


Figure 6: Simulations without CSI using $w_a : w_b : w_c = 1 : 0.5412 : 0.3827$.

Example 7 (BCMIM-3) : For this BCMIM, C_a, C_b, C_c, C_{bb} and C_{cc} are (8,1,8), (8,4,4), (8,4,4), (8,7,2) and (8,7,2) binary codes respectively. The code rate is 12/8 bits per signal point. In the 2-stage decoding, an 8-state trellis for the (16,11,4) binary code and a 1-state trellis for C_a are needed.

From simulation results shown in Fig. 4, for which perfect CSI is available, we see that all the three BCMIM have lower BER than the BCM-1 and BCM-2. Among the three BCMIM, BCMIM-2 which provides interblock memory between levels a and c is the best. This can be seen by

more detailedly examining the distance properties of these BCMIM, even though the MSD and MPD of these BCMIM are all identical to the MSD and MPD of BCM-2. In case that $u_r \neq u'_r$, we have $\delta_{Hr} = 6$ and $\Delta_{Pr}^2 = (2)^2 \times (0.586)^4$ and $N(\delta_{Hr}, \Delta_{Pr}^2) = 3584$ for BCMIM-1; $\delta_{Hr} = 6$ and $\Delta_{Pr}^2 = (4)^2 \times (0.586)^4$ and $N(\delta_{Hr}, \Delta_{Pr}^2) = 896$ for BCMIM-2; $\delta_{Hr} = 4$ and $\Delta_{Pr}^2 = (4)^2 \times (2)^2$ and $N(\delta_{Hr}, \Delta_{Pr}^2) = 1792$ for BCMIM-3. Since δ_{Hr} is more important than Δ_{Pr}^2 for high SNR, we can expect that BCMIM-2 which also has the smallest $N(\delta_{Hr}, \Delta_{Pr}^2)$ is the best. Although $N(\delta_{Hr}, \Delta_{Pr}^2)$ is larger and Δ_{Pr}^2 is smaller, BCMIM-1 is better than BCMIM-3 for high SNR, which is due to the larger δ_{Hr} .

From Fig. 5 and Fig. 6, we note that for all the three BCMIM, the usage of weight factors can improve the BER. We use BCMIM-2 as an example for explaining the reason. In the first stage of decoding for BCMIM-2, the SED between any two symbols in the first 8 positions is either 4 or 0 and the SED between any two symbols in the last 8 positions is either 0.586 or 0. In decoding this stage, the optimal ratio of weight factors is $(1/4)^{0.5} : (1/0.586)^{0.5} = 1 : 0.3827$. Compared with BCM-2, the coding rate of BCMIM-2 is increased from 9/8 to 12/8, and the error rate is lower. The price is that the decoding complexity for decoding BCMIM-2 is about twice of that for BCM-2. In the following, we slightly lower the coding rate of BCMIM-2 to more clearly show the advantage of BCMIM over BCM. **Example 8 (BCMIM-4) :** For this BCMIM, C_a, C_b, C_c, C_{aa} and C_{cc} are (8,1,8), (8,4,4), (8,4,4), (8,3,4) and (8,6,2) binary codes respectively. The code rate is 11/8 bits per signal point. In the 2-stage decoding, a 4-state trellis for the (16,7,4) binary code and a 4-state trellis for C_b are needed.

The decoding complexity of BCMIM-4 is almost the same as BCM-2. However, the coding rate is higher and the BER is lower.

References

- [1] G. Ungerboeck, "Channel coding with multi-level/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 1, pp. 55-67, Jan. 1982.
- [2] D. Divsalar and M.K. Simon, "The design of trellis coded MPSK for fading channels: Performance criteria," *IEEE Trans. Commun.*, vol. 36, pp. 1004-1012, Sept. 1988.
- [3] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inform. Theory*, vol. IT 23, pp. 371-376, May. 1977. 2
- [4] M.C. Lin and S.C. Ma, "A coded modulation scheme with interblock memory," *IEEE Trans. Commun.*, vol. 42, pp. 911-916, Feb./Mar./Apr. 1994.