The Analysis and the Performance Simulation of the Capacity of Bit-interleaved Coded Modulation System

[1] Hongwei ZHAO, [2] Mengmeng Shi, [2] Junfang Wang

[1] School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China
[2] Xi'an FengHuo Electronic Technology Co., Ltd., Xi'an 710075, China
[1] Tel.: 86-18092023949
[1] E-mail: Hongvi_zhao@126.com

Abstract: In this paper, the capacity of the BICM system over AWGN channels is first analyzed; the curves of BICM capacity versus SNR are also got by the Monte-Carlo simulations，and compared with the curves of the CM capacity. Based on the analysis results, we simulate the error performances of BICM system with LDPC codes. Simulation results show that the capacity of BICM system with LDPC codes is enormously influenced by the mapping methods. Given a certain modulation method, the BICM system can obtain about 2-3 dB gain with Gray mapping compared with Non-Gray mapping. Meanwhile, the simulation results also demonstrate the correctness of the theory analysis. *Copyright © 2014 IFSA Publishing, S. L.*

1. Introduction

Bit-interleaved coded modulation (BICM), which composed of coding, bit-interleaved, constellation mapping, was originally proposed by Zehavi in 1992 [1], and then, further research was carried out [2]. In the BICM system, coding and modulation is divided by bit-interleaved modules, so they can be designed individually, making the overall design of the system easy and flexible. In the receiving terminal, demodulator gets the information of bit based on the received data, and outputs the final result by de-interleaving and decoding [3, 4]. Coded modulation (CM) system is similar to the BICM. The trellis coded modulation (TCM), which was proposed by Ungerboeck in 1982, maximize the Euclidean distances between different ways by design jointly

the coding and modulation and set partition of the constellation points, which improved the performance. From the point of view of the capacity of channel, in fading channels, the performance of BICM is better than that of the trellis coded modulation (TCM) and symbol-interleaved coded modulation (SICM). However, in the additive white Gaussian noise (AWGN) channel, the performance of BICM is slightly worse than that of CM [1], the difference can be ignored under high signal-to-noise radio environment. Owe to the great robustness, BICM has become an industry standard in some fields of wireless communication, such as HSPA, IEEE802.11a/g and IEEE802.16e.

Low-density parity-check code (LDPC) is a kind of linear block code which has been widely used in the Date transmission and storage Systems. LDPC

was first proposed by Gallager in 1963 in his PHD dissertation. Nevertheless, due to the limit of hardware at that time, an available LDPC decoder is hard to be obtained, which leads to that little importance has been attached to LDPC for 30 years. In 1993, the finding of the Turbo coding aroused researchers' interest. And some researchers found that for regular LDPC, when adopting sum-product algorithm (SPA), the decoding performance is similar to that of Turbo code, and it even does better than Turbo in long code. The LDPC has low error floor and is able to realize parallel computing, meanwhile, it is easier to decoding than Turbo, making it be suitable for the hardware design and having the potential to decoding in high speed.

The paper analyzes the information between every bit and received numbers in mapping symbols firstly, then, the expression of the capacity of BICM is addressed, and finally, the Monte Carlo Simulation of BICM based on the LDPC is given and compares it with the BICM based on the RS under the same conditions. The simulation results show that the method of mapping of constellation deeply has a great effect on the performance of the BICM, the performance in Gray mapping is obviously better than that in Non-Gray mapping. Meanwhile, with the growth of the LDPC length, the performance of the BICM will get improved closed to the limit of the capacity of the BICM. At the same time, the results also demonstrate the validity of theory analysis.

2. BICM Model

The BICM model is just as shown in Fig. 1. Firstly, encoder codes the data flow $U = (u_0, u_1, u_2, \cdots, u_{k-1})$ produced by information source by using the linear block code whose parameter is (n, k), the length of the data flow U is k bits, and the output is $V = (v_0, v_1, v_2, \cdots v_{n-1})$, where $u_i, v_j \in \{0,1\}, 0 \le i \le k-1, 0 \le j \le n-1$.

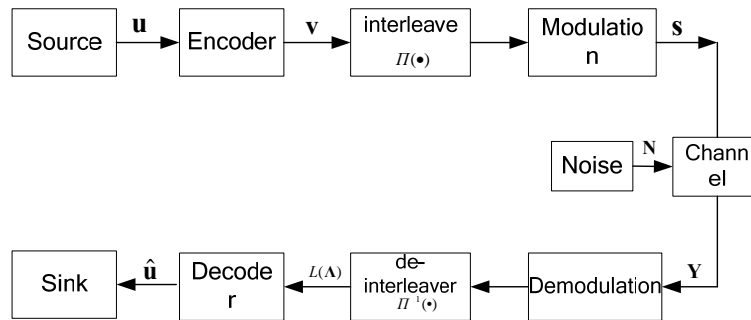**Fig. 1.** BICM model.

Then, interweave and modulate the V and get the output $S = (s_0, s_1, s_2, \cdots s_{q-1})$, where s_i and q are depended on the methods of modulation and constellation. Assume the constellation, in which the set of symbols is {A}, contains $|A| = 2^M$ symbols and each symbol can be expressed by M bits, that is, if $s_i = (b_i^0, b_i^1, b_i^2, \cdots b_i^{M-1})$, there will be $q = n/M$. Without loss of generality, taking 2-dimensional modulation into consideration, the Fig. 2 shows the constellations corresponding to the Gray mapping and the Non-Gray mapping under 8PSK and 16QAM modulation. In Fig. 2, the $\phi_1(x)$ and $\phi_2(x)$ denote a pair of orthogonal basis.

After the transmission of symbol sequence S in AWGN channel, the receiving terminal gets a complex sequence $Y = (y_0, y_1, y_2, \cdots y_{q-1})$, where the component $y_i = s_i + n_i$ is the superposed signal of the symbol s_i and noise, $n_i = n_i^R + jn_i^I$ is the component of the noise vector $N = (n_0, n_1, n_2, \cdots n_{q-1})$, n_i^R and n_i^I denote the real part and imaginary part of the complex noise respectively. And both of them are independent of each other and obey the Gauss distribution with 0 mean and $\sigma^2 = \dfrac{N_0}{2}$ variance. Demodulation module detects the maximum a posteriori (MAP) probability for the superposed signal, then gets the log-likelihood ratio (LLR) vector $L(\wedge)$ by de-interleaver, finally, the decoder complete the initializes and decoding. The output of the decoder is $\hat{U}$.

It should be noted that the interleaving module and de-interleaving module can be ignored when the LDPC is used as forward error correction code in BICM system.

3. Capacity Analysis of BICM

For simplicity, we focus on the capacity of the BICM under the conditions where the outputting of information source is binary and symbol equiprobability. That is, the priori probability $p(x=0) = p(x=1) = 0.5$ and for any ways of modulation, $p(s_k) = \dfrac{1}{2^M}, (0 \le k \le 2^M - 1)$.
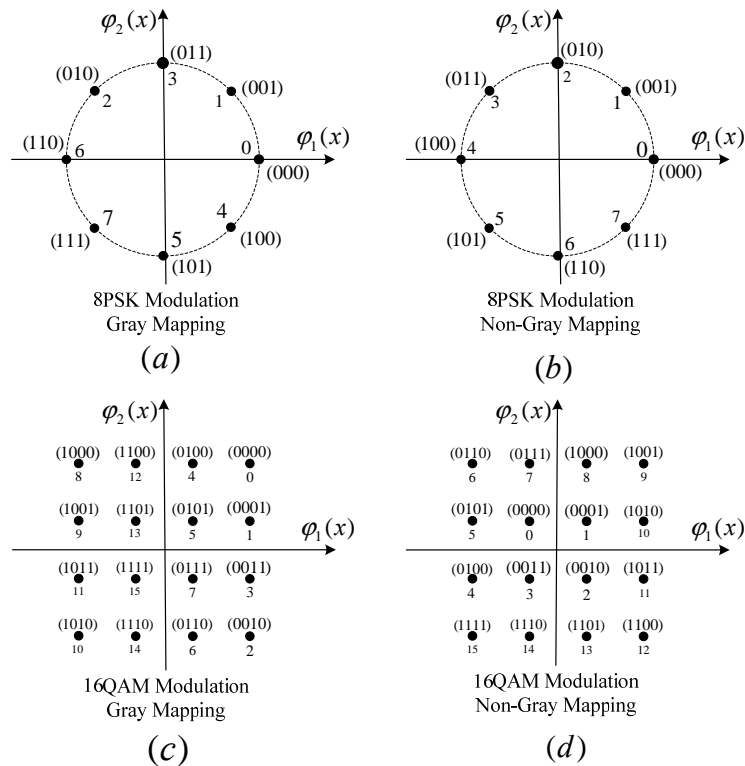
Fig. 2. Different methods of modulations and mapping.

In BICM system, the bit interleaving technique was employed in transmitting terminal and the maximum a posteriori (MAP) probability of detection was used in receiving terminal for the transformation between symbols and bit, so, it can be as the M parallel channels with independent binary input-continuous output. The Capacity of the BICM channels should be

$$I^{BICM}(X,Y) = \sum_{i=0}^{M-1} I(B_i,Y), \ B_i \in \{0,1\} , \quad (1)$$

where the symbol $I(B_i,Y)$ expresses the information between received value Y and the ith bit B_i in sending symbol X, the calculation process is as followed,

$$
\begin{aligned}
I(B_i,Y) &= H(B_i) - H(B_i \mid Y) \\
&= H(0.5) + \sum_{b_i,j} p(b_i,y_j) \log_2 p(b_i \mid y_j) , \\
&= 1 - E_{b_i,j}(\log_2 \frac{\sum_{s_k=\{A\}} p(y_j \mid s_k)}{\sum_{s_{k'}=\{A'\mid b_i=k\}} p(y_j \mid s_{k'})})
\end{aligned}
\quad (2)
$$

where $\{A' \mid b_i = k\}$ denotes the set of symbols which the value of the ith bit is k in constellation, $p(y_i \mid s_k)$ expresses the likelihood function, since the superposed noise is Complex Gaussian Noise, the likelihood function $p(y_i \mid s_k)$ can be expressed as follow,

$$
\begin{aligned}
p(y_j \mid s_k) &= \frac{1}{2\pi\sigma^2} e^{\frac{(y_i^R - s_k^R)^2 + (y_i^I - s_k^I)^2}{2\sigma^2}} \\
&= \frac{1}{2\pi\sigma^2} e^{\frac{\|y_i - s_k\|}{2\sigma^2}}
\end{aligned}
, \quad (3)
$$

where $\| y_i - s_k \|$ denotes the Euclidean distance between received complex signal y_i and the symbol s_k in constellation. Put (2) into (1) and we can get

$$I^{BICM}(X,Y) = M - \sum_{i=0}^{M-1} E_{b_i,j}(\log_2 \frac{\sum_{s_k=\{A\}} p(y_j \mid s_k)}{\sum_{s_{k'}=\{A'\mid b_i=k\}} p(y_j \mid s_{k'})}), \quad (4)$$

From the expression of the BICM capacity, we can find that under the premise of giving the way of modulation, the mapping relationship between bit and symbol has greatly important influence on the result. Meanwhile, from (2) we also can find that the information between Y and different bit $B_i (0 \le i \le M-1)$ in mapping symbols is different. Based on it, the BICM system is able to adopt different modulation method to carry out the unequal error protection (UEP). Through putting important bit information into the position where the $I(B_i,Y)$ is larger, it obtain reliable information with high probability after the demodulation. The Fig. 3(a) shows the relationship between the capacity $I^{BICM}(X,Y)$ of BICM system and SNR($\frac{1}{\sigma^2}$)

corresponding to the Gray mapping and the Non-Gray mapping (Fig. 2) under the 8PSK and 16QAM modulation. For comparison, Fig. 3(a) shows the capacity of the CM under 8PSK and 16QAM, and the expression of the CM capacity is

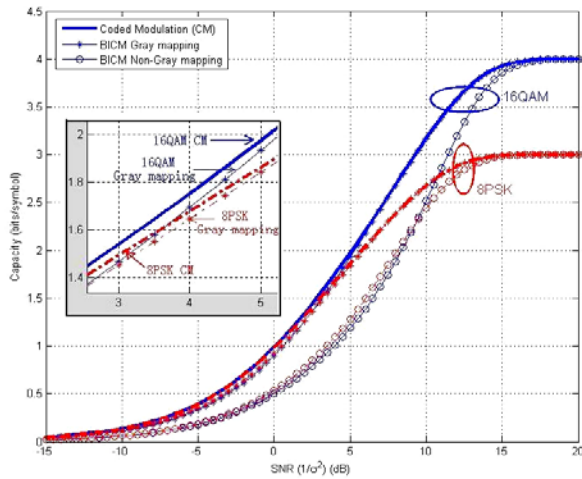$$I^{CM}(X,Y) = \boldsymbol{E}_{k,j}(\log_2 \frac{p(y_j \mid s_k)}{\sum\limits_{s_k=\{A\}} p(s_k)p(y_j \mid s_k)}) , \quad (5)$$

From Fig. 3, we can see the following points.

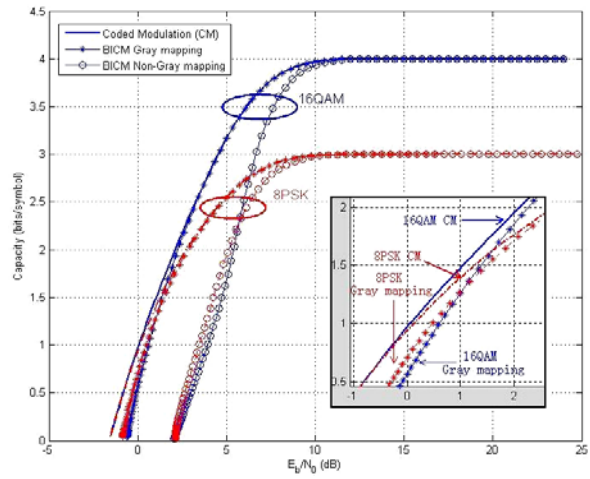1) Regardless of 8PSK or 16QAM, the capacity of BICM corresponding to the Gray mapping is obviously larger than Non- Gray mapping under the same SNR;

2) The capacity of BICM and CM is almost the same when using the Gray mapping under the circumstance with high SNR; but when the SNR is low, the capacity of the BICM is slightly smaller than that of CM.

In order to reflect clearly the effect of the mapping methods on the BICM capacity, the relationship between $I^{BICM}(X,Y)$ and E_b/N_0 of the BICM is provided as Fig. 3(b). In the next sections about performance simulation, we will choose E_b/N_0 will be chose as the parameter of simulation.

(a) BICM capacity vs. SNR   (b) BICM capacity vs. E_b/N_0

Fig. 3. The channel capacity curves.

4. Expression of the LDPC and Decoding Algorithm

Here we take the binary LDPC code as an example to introduce the expression method and the decoding algorithm.

4.1. Expression of the LDPC

LDPC can be expressed by a matrix. Let $\boldsymbol{H} = (h_{i,j})_{m \times n}$ be a sparse $m \times n$ check matrix. When $\boldsymbol{V}$ satisfies $\boldsymbol{HV}^T = \boldsymbol{0}$, $\boldsymbol{V}$ is a code of LDP defined in check matrix $\boldsymbol{H}$. Wherein, $V \in \mathrm{F}_2^n$.

Define the line weight (or column weight) of the i-th line i(or j-th column) in matrix H as the number of its non-zero elements, and call the LDPC code with constant line (column) weight for $\boldsymbol{H}$ as regular LDPC mode, otherwise, call them non-regular LDPC mode. For regular LDPC mode, the code rate

$$R \geq 1 - \frac{m}{n} .$$

A check matrix of LDPC corresponds to a Tanner graph [9]. Tanner graph is a bipartite graph, which contains two types of nodes: variable nodes and check nodes. The number of variable nodes is n, which is corresponding to n bit codes. There are m check nodes which are corresponding to the m check equation. If the elements $h_{i,j}(0 \leq i < m, 0 \leq j < n)$ in matrix $\boldsymbol{H}$ are not equal to 0, an edge is introduced between i-th check node and j-th variable point. Fig. 4 shows a check matrix of LDPC and its corresponding Tanner graph.

The use of Tanner graph can describe iterative decoding algorithm more clearly, including the information processing flow and the edge information flow of different types of node.

4.2. SPA Decoding Algorithm of LDPC Code

Let $u_{i,j}^{(l)}$ be the information transmitted from check node c_i to variable node v_j in l-th iteration. Let $v_{j,i}^{(l)}$ represent the information transmitted from variable node v_j to check node c_i in l-th iteration.
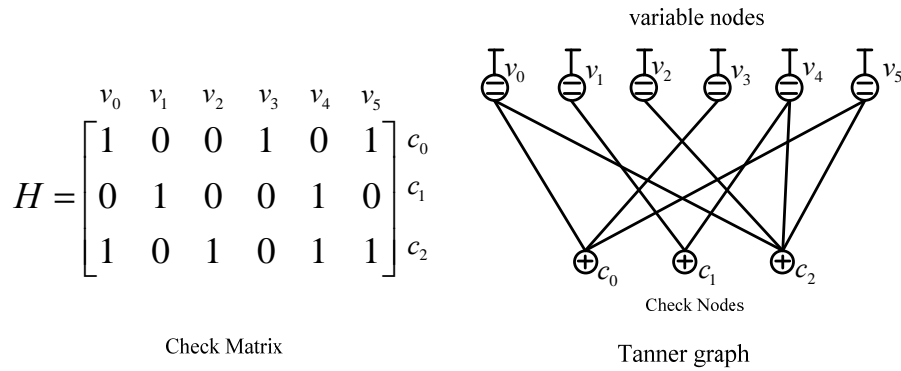
Fig. 4. Check matrix of LDPC and the corresponding Tanner graph.

Demodulation

Using the received complex-value $y_i (0 \le i \le q-1)$, we can calculate the log-likelihood ratio information corresponding to each bit in the symbol.

$$\Lambda_i^j = \ln \frac{\sum_{s_k \in \{A_0 | b_i^j = 0\}} p(y_i | s_k)}{\sum_{s_{k'} \in \{A_1 | b_i^j = 1\}} p(y_i | s_{k'})}, \quad (0 \le j \le M\text{-}1), \quad (6)$$

where symbol $\{A_k | b_i^j = k\}$ denotes the set of the symbol that the value of *i*-th symbol, *j*-th bit is k in constellation. The likelihood function $p(y_i | s_k)$ can be computed according to equation (3). Based on equation (6), the initial log-likelihood ratio information $L(\Lambda)$ can be computed according to equation (6) for all bit, and $L(\Lambda)$ is looked as the initial information $v_j^{(0)} (0 \le j \le n-1)$ of decoder.

Information process of check node

Assuming that check node c_i received the information come from its neighboring node in the *l*-th iteration. The information $u_{i,j}^{(l)}$ transmitted from check node c_i to variable node v_j can be calculated as follow:

$$u_{i,j}^{(l)} = 2\tanh^{-1}(\prod_{j' \in N_{i/j}} \tanh(\frac{1}{2} v_{j',i}^{(l)})), \quad (7)$$

where symbol $N_{i/j}$ denotes the variable nodes corresponding to the check node c_i besides node v_j.

The information process of variable node

In the *l*-th iteration, the information $v_{j,i}^{(l+1)}$ transmitted from variable node v_j to check node c_i can be calculated as follow:

$$v_{j,i}^{(l+1)} = v_j^{(0)} + \sum_{i' \in M_{j/i}} u_{i',j}^{(l)}, \quad (8)$$

where symbol $M_{j/i}$ denotes the check nodes connected to the variable node v_j besides c_i.

When every iteration comes to the end, in order to get the estimation of V, decoder need to make hard-decision. The methods of information update and hard-decision are as follow:

$$L(v_j) = v_j^{(0)} + \sum_{i \in M_j} u_{i,j}^{(l)}, \quad (9)$$

$$\hat{v}_j = \begin{cases} 0 & \text{if } L(v_j) \ge 0 \\ 1 & \text{else} \end{cases}, \quad (10)$$

According to the above information processing, we make the summary for SPA algorithm and describe as follow, where J is the max numbers of iteration in the process of decoding.

Sum-product algorithm

• **Initialization**. For all variable nodes, calculate the initial LLR information $v_j^{(0)}$ according to equation (6), meanwhile, let $v_{j,i}^{(0)} = v_j^{(0)}$, and the number of iteration is set to 0.

• **Iteration**. When $l < J$, carry out the steps as follows:

1. **Check nodes**. For all check nodes, calculate the information $u_{i,j}^{(l)}$ that transmitted to variable nodes from check nodes according to equation (7).

2. **Variable nodes**. For all variable nodes, calculate the information $v_{j,i}^{(l)}$ that transmitted from variable nodes to check nodes according to equation (8).

3. **Decision**: for all variable nodes, calculate $L(v_j)$ according to equation (9), and make hard-decision according to equation (10);

4. If $H\hat{V}^T = 0$, let $l = J$, or else let $l = l+1$.

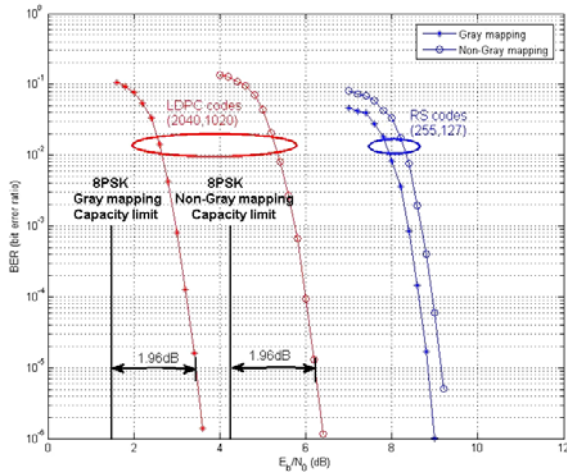• **Output**. Output $\hat{V}$ as the result of decoding.

5. Simulations

In our simulations, the performance of the BICM system show as Fig. 1 is investigated. When the LDPC code is used as forward error correction (FEC), the interlaver / deinterlaver can be ignored. The decoding algorithm for LDPC adopts the SPA
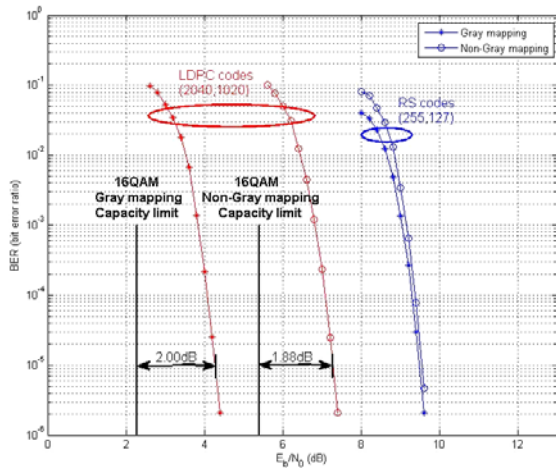
algorithm presented in section 3.2, where the maximum number of iterations is 50.

Simulation 1.

Select LDPC codes of IEEE 802.16e standard as FEC code whose rate is 0.5 and length is 2040. Modulation and mapping method are shown as Fig. 2, which are 8 PSK/Gray mapping, 8PSK/Non-Gray mapping, 16 QAM/Gray mapping and 16 QAM/Non-Gray mapping respectively. The stimulation results are shown in Fig. 5. Meanwhile, For comparison easily, a performance curve of RS code is simulated, and the equivalent length of the RS code is 2040, defining in $GF(2^8)$.

(a) 8PSK

(b) 16QAM

Fig. 5. Performance curve of BICM based on (2040, 1020) LDPC code of IEEE 802.16e.

From Fig.5 we can see the following points.
When the modulation is 8PSK.

1) The effect of mapping method on LDPC code is far more than that on RS code. For example, take the bit error rate 10^{-5} as judge criterion, for Gray mapping and Non-Gray mapping, the SNR of LDPC code should be 3.440 dB and 6.221 dB respectively, which the difference is 2.781 dB; while the SNR of

RS code are 8.837 dB and 9.145 dB respectively, the difference is 0.308 dB. That difference is that the RS code are defined in finite fields, the decoding process is to symbol, while the decoding process of LDPC code is to bits.

2) Under the same condition, LDPC code can reduce the distance to the theoretical capacity limit. Take the bit error rate 10^{-5} as criterion, the distance to theoretical capacity limit of LDPC code and RS code are 1.960 dB and 7.355 dB in Gray mapping, and there are 1.960 dB and 4.890 dB in Non-Gray mapping.

When the modulation method is 16 QAM, the conclusion is 8 PSK.

Simulation 2.

In this simulation, we choose LDPC codes of DVB S2.0 standard as FEC code whose rate is 0.5, and code length is 6400. In order to evaluate the effect of different code length on the BICM performance, we draw the performance curve under 8PSK modulation and Gray/Non-Gray mapping. The simulation results are shown in Fig. 6. In order to facilitate comparison, the LDPC code performance curve based on IEEE 802.16e in simulation 1 is also drew in Fig. 6.
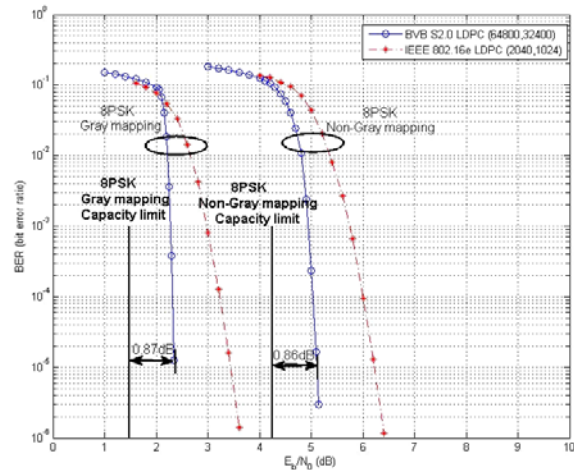

Fig. 6. Performance curve of BICM based on (64800, 32400) LDPC code of DVB S2.0.

From Fig. 6 we can see that with the increase of LDPC code length, the performance of BICM system is approaching the capacity limit. For example, take bit error rate 10^{-5} as criterion, the LDPC code of DVB S2.0 required the SNR of 2.350 dB and 5.115 dB respectively under Gray and non-Gray mapping. Compared to the LDPC code of IEEE 802,16e, the gain is 1.09 dB and 1.10 dB.

6. Conclusions

In this paper, we analyzed the capacity that BICM system in AWGN channels firstly, then, by using the Monte Carlo simulation, we get the relationship curve between channel capacity and SNR under different modulation and Gray/Non-Gray mapping,

furthermore, we compared the channel capacity to CM system under the same conditions. Finally, on the basis of the performance simulation, the performance of the BER is evaluated. Meanwhile, LDPC code and RS code is regard as FEC respectively, and the difference between them is compared. Simulation results show that the mapping method of constellation will influence the performance of BICM system which based on LDPC code in large extent, the performance with Gray mapping is better than that with Non-Gray mapping clearly, and with the increase of LDPC code length, the performance of BICM system is approaching the capacity limit.

Acknowledgements

References

[1]. E. Zehavi, 8-PSK trellis codes for a Rayleigh channel, *IEEE Transaction on Communication*, Vol. 40, Issue 3, 1992, pp. 873–884.

[2]. G. Caire, G. Taricco, and E. Biglieri, Bit-interleaved coded modulation, *IEEE Transaction on Information Theory*, Vol. 44, Issue 3, 1998, pp. 927–946.

[3]. G. Ungerboeck, Channel coding with multilevel/phase signals, *IEEE Transaction on Information Theory*, Vol. 28, Issue 1, 1982, pp. 55-67.

[4]. R. G. Gallager, Low-density parity-check codes, *MIT Press*, Cambridge, Mass, 1963.

[5]. C. Berrou, A. Glavieux, P. Thitimajshima, Near Shannon limit error-correcting coding and decoding: Turbo codes, in *Proceedings of the IEEE International Conference on Communications*, 1993, pp. 1261-1271.

[6]. R. C. Bose, D. K. Ray-Chaudhuri, On a class of error correcting binary group codes, *Information and Control*, Vol. 3, 1960, pp. 68-79.

[7]. R. M. Tanner, A recursive approach to low complexity codes, *IEEE Transaction on Information Theory*, Vol. 27, Issue 5, 1981, pp. 533-547.

————————————————