Mapping Optimization for Space–Time Bit-Interleaved Coded Modulation With Iterative Decoding

Wookbong Lee, Jungho Cho, Chang-Kyung Sung, Hwangjun Song, and Inkyu Lee, *Senior Member, IEEE*

***Abstract*—For space–time bit-interleaved coded modulation (ST-BICM) systems with iterative decoding, the overall performance is affected by the chosen mapping. In bit-error rate (BER) curves, one mapping reaches an error floor (EF) at a low signal-to-noise ratio (SNR), while other mappings result in a lower EF at a higher SNR. The constellation mappings are divided into groups where each group exhibits a distinctive BER curve. We show that the convergence abscissa of the system depends on the average total bit errors and the harmonic mean of the minimum squared Euclidean distance. In this letter, we characterize all mapping groups for ST-BICM with 8-phase-shift keying and present the optimal selection for each mapping group over independent fading channels.**

***Index Terms*—Bit-interleaved coded modulation (BICM), coded modulation, iterative decoding, space–time (ST) code, symbol labeling.**

I. Introduction

SPACE–TIME (ST) coding techniques significantly improve transmission efficiency in radio channels by using multiple transmit and/or receive antennas and coordination of the signaling over these antennas. Bit-interleaved coded modulation with iterative decoding (BICM-ID) gives good diversity gains and Euclidean distance property with higher order modulation schemes using well-known binary convolutional codes [1]. Tonello [2] combined these two schemes as ST-BICM with Gray mapping. A reduced-complexity receiver for the ST-BICM was studied in [14]. There exist other approaches which combine modulation and coding using non-Gray mapping for bit-interleaved ST-coded modulation with iterative decoding [3] and low-density parity-check (LDPC)-coded modulation [4].

In this letter, we will first show that the asymptotic bit-error rate (BER) performance of ST-BICM systems is equivalent to the single-antenna BICM-ID case even when the throughput of ST-BICM is higher. This indicates that the total data rate of ST-BICM can be increased linearly with the number of transmit antennas while maintaining the same error probability as BICM-ID. This demonstrates the superior performance of ST-BICM. Motivated by the fact that the ST-BICM performance can be characterized by BICM-ID with the same code and mapping, we optimize the ST-BICM mappings based on the BICM-ID parameters.

While the error floor (EF) bound [1] offers a good match for an EF region, a heuristic analysis is still needed to study BER curve behaviors of BICM-ID for the waterfall region. Thus, we derive design considerations based on both analytical approaches and actual simulations. In this letter, rather than attempting to find a mapping which achieves the lowest EF as examined in [5], we characterize all mapping groups for ST-BICM/BICM-ID with 8-phase-shift keying (PSK), and present the optimal selection for each mapping group over channels with independent fadings. As a result, we provide design guidelines such that a proper mapping can be selected depending on the system requirement.

This letter is organized as follows. In Section II, we review ST-BICM systems, and we show that the asymptotic performance of both BICM-ID and ST-BICM is the same in Section III. In Section IV, a mapping design criterion over independent fading channels and the optimal mapping for ST-BICM/BICM-ID are presented. Finally, in Section V, extensive simulation results are provided to verify our analysis.

II. Review of the ST-BICM System

Fig. 1 shows the transmitter and receiver structure of the ST-BICM system. N_u data bits $\underline{u}$ are encoded by a convolutional encoder, and then N_c coded bits $\underline{c}$ are interleaved by a random interleaver π. The resulting code rate is $R_c = N_u/N_c$. Then, the interleaved bits $\underline{v}$ are serial-to-parallel (S/P) converted into N_t transmit antennas as $[\underline{v}^1, \underline{v}^2, \ldots, \underline{v}^{N_t}]$, where $\underline{v}^i$ is a vector of length $N_v = N_c/N_t$. At time t, each stream is grouped into m bit subsequences to form a channel symbol $x_t^i = \mu([v_t^{i,1}, v_t^{i,2}, \ldots, v_t^{i,m}])$ for $i = 1, \ldots, N_t$, chosen from the M-ary signal constellation χ according to the labeling map

Paper approved by R. W. Heath, the Editor for MIMO Techniques of the IEEE Communications Society. Manuscript received April 9, 2005; revised February 12, 2006 and June 26, 2006. This work was supported in part by the Ministry of Information and Communication (MIC), Korea, under the Information Technology Research Center (ITRC) support program, supervised by the Institute of Information Technology Assessment (IITA) , and in part by the Basic Research Program of the Korea Science and Engineering Foundation under Grant R01-2006-000-11112-0. This paper was presented in part at the IEEE Vehicular Technology Conference, September 2005.

W. Lee was with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea. He is now with the Mobile Communication Technology Research Laboratory, LG Electronics Institute of Technology, Kyunggi-do 431-080, Korea (e-mail: wbong@lge.com).

J. Cho was with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea. He is now with the Agency for Defense Development, Daejun 305-152, Korea (e-mail: jungho@add.re.kr).

C.-K. Sung was with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea. He is now with Samsung Electronics, Kyunggi-do 443-742, Korea (e-mail: cksung@samsung.com).

H. Song is with the Department of Computer Science and Engineering, Pohang University of Science and Technology (Postech), Pohang 798–784, Korea (e-mail: hwangjun@postech.ac.kr).

I. Lee is with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea (e-mail: inkyu@korea.ac.kr).

Digital Object Identifier 10.1109/TCOMM.2007.892439

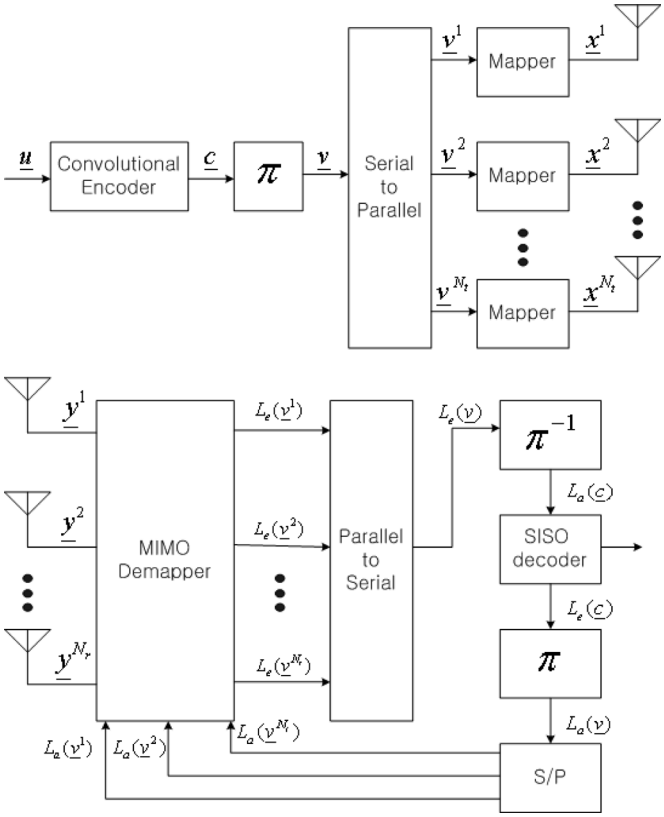

Fig. 1. ST-BICM transmitter and receiver structure.

μ. Then, assuming flat fading channels, N_r received signals at time t are formulated as

$$\begin{bmatrix} y_t^1 \\ \vdots \\ y_t^{N_r} \end{bmatrix} = \begin{bmatrix} h_t^{1,1} & \dots & h_t^{N_t,1} \\ \vdots & \ddots & \vdots \\ h_t^{1,N_r} & \dots & h_t^{N_t,N_r} \end{bmatrix} \begin{bmatrix} x_t^1 \\ \vdots \\ x_t^{N_t} \end{bmatrix} + \begin{bmatrix} n_t^1 \\ \vdots \\ n_t^{N_r} \end{bmatrix}$$

or more compactly, $\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{n}_t$, where y_t^j is the jth received symbol, $h_t^{i,j}$ denotes the channel coefficients from the ith transmit antenna to the jth receive antenna, independent complex Gaussian with zero mean (Rayleigh fading), and n_t^j is modeled as an additive white complex Gaussian noise with zero mean and variance σ^2.

At the receiver shown in Fig. 1, the demapper extracts the extrinsic log-likelihood ratio (LLR) values of the estimated symbol sequence by the maximum *a posteriori* (MAP) rule as

$$L_e\left(v_t^{i,n}\right) = \log \frac{\sum_{\mathbf{x}_t \in \chi_0^{i,n}} p(\mathbf{x}_t, \mathbf{y}_t, \mathbf{H}_t)}{\sum_{\mathbf{x}_t \in \chi_1^{i,n}} p(\mathbf{x}_t, \mathbf{y}_t, \mathbf{H}_t)} \tag{1}$$

where $\chi_b^{i,n}$ denotes the signal subset $\{\mu([v_t^{i,1}, v_t^{i,2}, \dots, v_t^{i,m}]) \mid v_t^{i,n} = b\}$ for any t.

Then, the extrinsic LLR values are parallel-to-serial (P/S) converted as $L_e(\underline{v})$. After deinterleaving, these values are applied as *a priori* LLR values for the MAP decoder. During the iterative decoding process, the extrinsic LLR values $L_e(\underline{c})$ out of the MAP decoder are interleaved as $L_a(\underline{v})$, and fed back to the MAP demapper as *a priori* LLR values $[L_a(\underline{v}^1), L_a(\underline{v}^2), \dots, L_a(\underline{v}^{N_t})]$ for the next iteration.

The joint probability density function in (1) is related to

$$p(\mathbf{x}_t, \mathbf{y}_t, \mathbf{H}_t) \propto \exp\left(-\frac{1}{\sigma^2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_t\|^2 + \frac{1}{2}\mathbf{v}_t^T\mathbf{L}_t\right)$$

where $\mathbf{v}_t$ and $\mathbf{L}_t$ are column vectors of length $N_t \cdot m$, comprised of $v_t^{i,n}$ and $L_a(v_t^{i,n})$, respectively [6], [7]. At the first iteration, the *a priori* LLR values $L_a(\underline{v})$ are set to zero.

III. Performance Analysis for ST-BICM

At high signal-to-noise ratios (SNRs), the asymptotic performance of ST-BICM over independent Rayleigh fading channel is given as

$$\log_{10} P_b \simeq -\frac{N_r \cdot d_H}{10}\left[\left(d_h^2\right)_{\text{dB}} + \left(\frac{1}{\sigma^2}\right)_{\text{dB}}\right] + C \tag{2}$$

where d_H denotes the free Hamming distance of the code, d_h^2 is defined as the harmonic mean of the minimum squared Euclidean distance, and C represents a constant number which is independent of the choice of the mapping or the channel code [1], [8] (see the Appendix for proof). Here $N_r \cdot d_H$ determines the slope of error probability, and the harmonic mean of the minimum squared Euclidean distance d_h^2 affects the horizontal shift of error probability curves where d_h^2 is

$$d_h^2 = \left(\frac{1}{m \cdot 2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{x \in \chi_b^i}\left(\frac{1}{|x-z|^2}\right)^{N_r}\right)^{-1/N_r}. \tag{3}$$

Here z denotes an error symbol $z \in \chi_{\bar{b}}^i$, where $\bar{b}$ denotes the complement of b. When computing d_h^2, z should be determined depending on the given system assumption. For example, for the noniterative BICM case, z corresponds to the nearest neighbor of x with the ith bit $\bar{b}$, $z = \hat{z}$ [8]. In contrast, with the ideal feedback assumption for each $x \in \chi_b^i$, z is set to $\tilde{z} \in \chi_{\bar{b}}^i$, whose label has the same binary bit values as those of x except at the ith bit position [1].

We denote d_h^2(before) and d_h^2(after) as the harmonic mean of the minimum squared Euclidean distance before and after feedback, which can be obtained by plugging $\hat{z}$ and $\tilde{z}$ to z in (3), respectively [1], [8]. For example, d_h^2(before) accounts for the coding gain for noniterative systems. Thus, at low SNRs where the performance is close to the noniterative BICM case, a mapping with a large d_h^2(before) exhibits a better performance, while at high SNRs where the ideal feedback can be assumed, one with a large d_h^2(after) performs better for the EF region.

Note that (2) is independent of the number of transmit antennas N_t. Most importantly, this equation is the same as the BICM-ID case [8], except for the number of receive antennas N_r. Here the overall data rate of ST-BICM is N_t times larger than the single-antenna BICM-ID case. Thus (2) indicates that the asymptotic performance of ST-BICM is the same as BICM-ID, even when the spectral efficiency is increased by N_t. In other words, the total throughput of ST-BICM can be increased linearly with the number of transmit antennas while maintaining the same performance. This demonstrates an

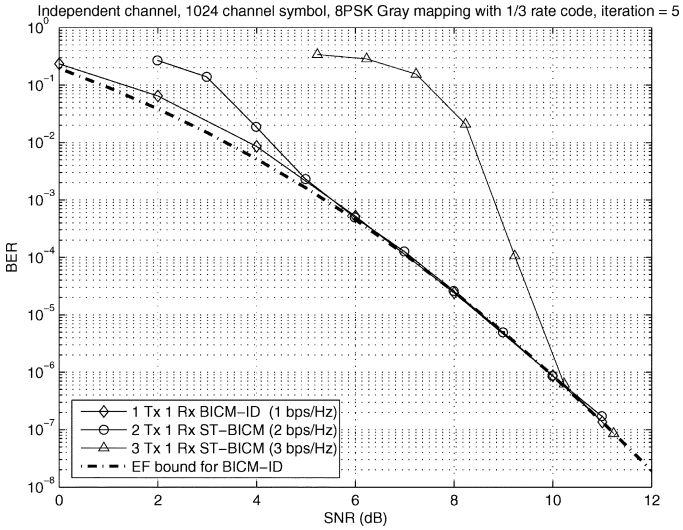

Fig. 2. ST-BICM/BICM-ID with various numbers of transmit antennas.

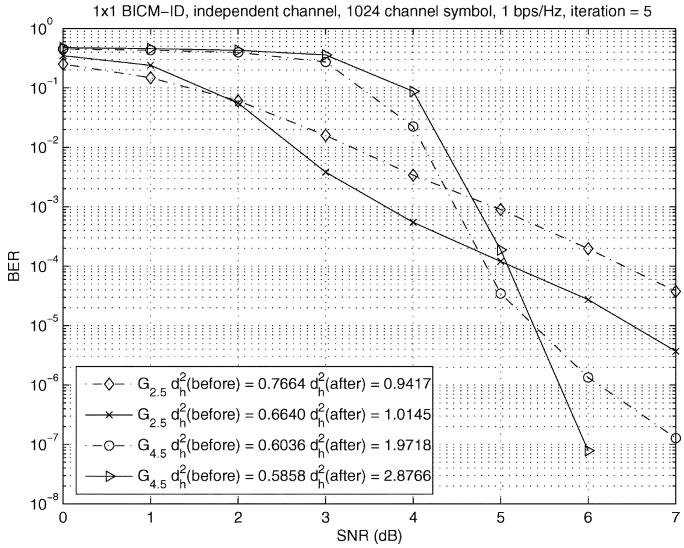

Fig. 3. Two different mapping groups $G_{2.5}$ and $G_{4.5}$.

outstanding characteristic which makes ST-BICM an attractive solution.

Fig. 2 confirms our analysis by comparing the performance of ST-BICM/BICM-ID with different numbers of transmit antennas combined with Gray mapping and a 16-state rate-1/3 convolutional code [9] without transmit power normalization. The figure demonstrates that BER curves exactly match with the BICM-ID EF bound with 1 b/s/Hz at high SNRs, while the spectral efficiency is increased by N_t. Note that the exact ST-BICM EF bound can be directly evaluated from the above equations as in [8], which is exactly the same as the BICM-ID EF bound. Because of the fact that the ST-BICM performance is equivalent to the BICM-ID performance at high SNRs, we can resort to the optimum parameters of BICM-ID when optimizing the ST-BICM in the following section.

IV. Mapping Optimization

As shown in Section III, the system performance for BICM-ID and ST-BICM is asymptotically equivalent. This motivates us to optimize the ST-BICM mappings based on the BICM-ID parameters. Thus, from now on, we first focus on the optimization of BICM-ID parameters. Then, those optimized parameters will be applied to the ST-BICM later in the simulation sction.

In general, performance curves for iterative decoding schemes are divided into three regions: nonconvergence region, waterfall region, and EF region [10]. Gray mapping is shown to be optimum for the nonconvergence region [8], while a semi-set partitioning (SSP) mapping is optimum for the EF region for BICM-ID [1], [5]. At the waterfall region, BER curves lie between the first-pass performance curve and the EF performance curve. The first-pass performance can be optimized by maximizing $d_h^2(\text{before})$, while the EF performance can be optimized by maximizing $d_h^2(\text{after})$ [1].

In this section, rather than attempting to find the best mappings for any specific region, we characterize all possible mappings for 8PSK. Noting that there exist $4 \cdot 6!$ different mappings for 8PSK, we classify each mapping into distinct groups based on parameters such as d_h^2. Another parameter which can be used for characterizing each group is the average total bit errors N_b, defined in [11] as

$$N_b = \sum_{i=0}^{M-1} p_x(i) \sum_{j=1}^{N_i} n_b(i,j)$$

where $p_x(i)$ denotes the probability of the ith constellation symbol $x(i)$, N_i is the number of neighboring constellation points of $x(i)$, and $n_b(i,j)$ represents the number of bit errors when $x(i)$ is erroneously detected as $x(j)$. In the waterfall region, the overall performance is dominated by the demapper performance, which can be approximated by $P_{\text{demap}} \approx N_b P_e$ [11] where P_e denotes the symbol-error probability (SEP) as a function of the minimum Euclidean distance of the constellation and SNR. Therefore, N_b should be minimized when identifying a good mapper for the waterfall region. Note that N_b does not affect the performance for the EF region, since with the ideal feedback assumption, an erroneously detected symbol has only one bit error.

For 8PSK, the Gray mapping has $N_b = 2$ and shows the best first-pass performance, and the SSP mapping exhibits the best EF performance with $N_b = 4.5$. To optimize the waterfall region, we classify all possible constellation mappings with respect to N_b. We refer to G_{N_b} as a group of mappings with N_b. Note that mappings which belong to the same group G_{N_b} could exhibit different performance behaviors among themselves, since they have diverse values of $d_h^2(\text{after})$ and $d_h^2(\text{before})$. Fig. 3 shows BER curves for two different mapping groups, $G_{2.5}$ and $G_{4.5}$. The solid lines represent the performance curve for the mapping with the maximum $d_h^2(\text{after})$, while the dashed lines indicate that for the mapping with the maximum $d_h^2(\text{before})$ within each group. From this plot, we observe that $d_h^2(\text{after})$ plays a more important role in the BER curves than $d_h^2(\text{before})$.

Thus, we provide the mapping optimization process for ST-BICM/BICM-ID as follows.

TABLE I
HARMONIC MEANS OF THE MINIMUM SQUARED EUCLIDEAN DISTANCE AND THE ITERATION GAIN FOR MAPPING GROUPS ($N_r = 1$)

	μ	d_{h}^2(before)	d_{h}^2(after)	Iteration Gain (dB)
small	$G_{2.0}$	0.7664	0.8093	0.2361
gain	$G_{2.5}$	0.6640	1.0145	1.8406
groups	$G_{3.0}$	0.6640	1.2632	2.7926
large	$G_{3.5}$	0.5858	1.9444	5.2105
gain	$G_{4.0}$	0.5858	2.3204	5.9782
groups	$G_{4.5}$	0.5858	2.8766	6.9114

- For all possible constellation mappings, compute d_h^2(before), d_h^2(after), and N_b.
- Classify mappings into groups with respect to N_b.
- Identify mappings with the maximum d_h^2(after) within each group.
- Among the mappings with the maximum d_h^2(after) within a group, choose one of the mappings with the maximum d_h^2(before). (There may be several mappings with the maximum d_h^2.)

Table I lists the mapping groups of the 8PSK modulation for one receive antenna based on the above rule, and Fig. 4 shows one of the optimal selections for each mapping group. Denoting the iteration gain as d_h^2(after)/d_h^2(before), which accounts for the asymptotic gain due to iterative decoding, we can divide mapping groups into two groups with respect to the iteration gain: the small gain groups and the large gain groups. As explained before, a mapping with small N_b and large d_h^2(before) performs better at the nonconvergence region and the waterfall region, while a large d_h^2(after) results in a good performance for the EF region. It is clear from Table I that the small gain groups have better N_b and d_h^2(before), whereas the large gain groups have better d_h^2(after). Therefore, it is expected that the small gain groups perform better for the waterfall region, while the large gain groups have better performance for the EF region. Because of this observation, we also expect that the performance curves of the small gain groups and the large gain groups will cross over with each other, and this will be confirmed by the simulation results in Section V.

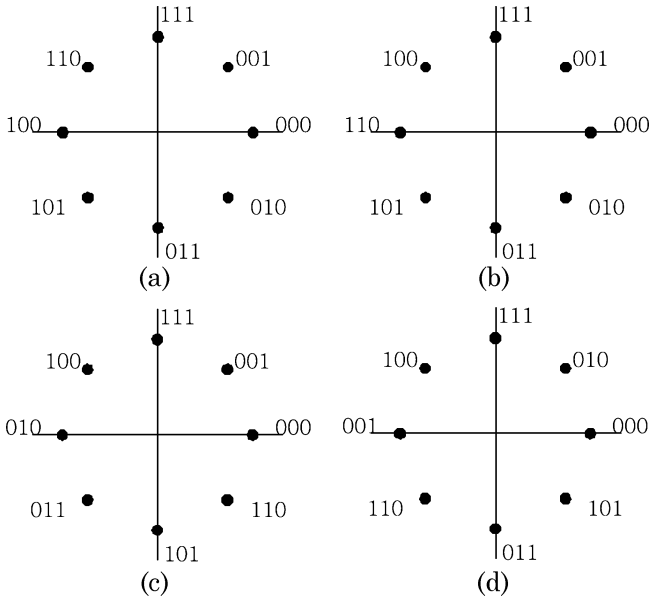

Fig. 4. Optimal mapping. (a) $G_{2.5}$. (b) $G_{3.0}$. (c) $G_{3.5}$. (d) $G_{4.0}$.

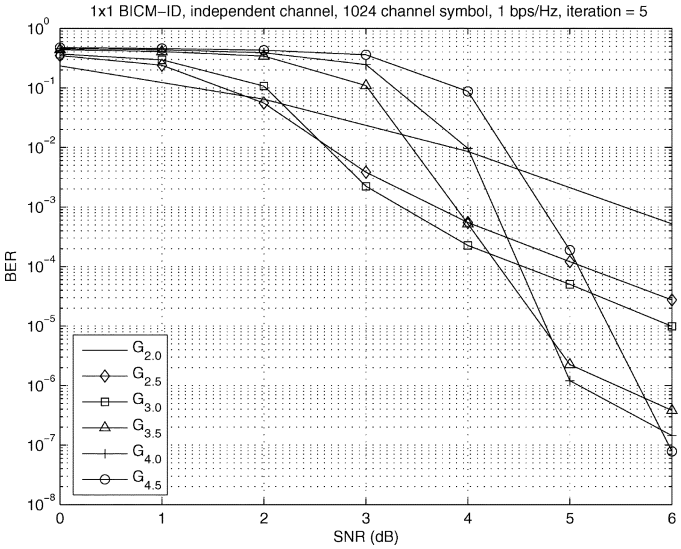

Fig. 5. 1×1 BICM BER curve for 1 b/s/Hz over independent Rayleigh fadings.

V. SIMULATION RESULTS

In this section, we employ an iterative decoding algorithm with five iterations, and the symbol length is set to 1024. The channel response between the transmitter and the receiver is assumed to be independent Rayleigh fading. A punctured convolutional code with constraint length 5 [9] and a random interleaver are used in this simulation with 8PSK.

In Fig. 5, the simulation results for the 1×1 BICM 1 b/s/Hz system with a rate-1/3 convolutional code are presented. For low-to-moderate SNRs, the small gain groups $(G_{2.0}, G_{2.5}, G_{3.0})$ perform better than the large gain groups $(G_{3.5}, G_{4.0}, G_{4.5})$. As SNR increases, the BERs of the large gain groups decrease rapidly, and eventually outperform the small gain groups because of their high iteration gain. This is exactly what we expect from the parameter analysis in Table I.

Fig. 6 shows BER curves of the 2×1 ST-BICM system for 2 b/s/Hz with a rate-1/3 code. This plot exhibits a similar pattern with the 1×1 BICM case. Compared with the BICM-ID system, the performance improvement from iterative decoding becomes large even for the small gain groups for the ST-BICM system, especially when $N_t > N_r$, since the iterative decoding process compensates for the signal-processing loss in the demapper block. This makes the crossover points of the ST-BICM lower than those of the BICM-ID system with the same code rate. Thus, the crossover points are sometimes not observed at the SNR operating range. For example, in Fig. 6, some crossover points occur near BER $= 10^{-8}$. In this case, it is much better to employ mappings in the small gain groups.

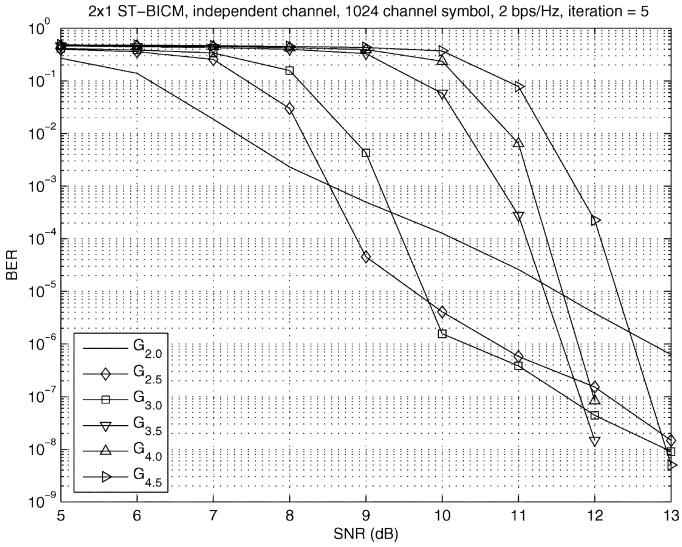

Fig. 6. 2×1 ST-BICM BER curve for 2 bps/Hz over independent Rayleigh fading channel.

Another important observation which can be made in the plots is that our analysis on the asymptotic performance in Section III is accurate. Comparing Fig. 5 with Fig. 6, we can see that the BER curves of ST-BICM at the EF region are 3 dB away from the BICM-ID curves due to transmit power normalization. These results exactly confirm our analysis in Section III.

VI. Conclusion

We have shown that the total data rate of ST-BICM can be increased by N_t compared with the single-antenna BICM-ID case while maintaining the same asymptotic performance. Then, we have characterized all mappings and presented the optimal mapping for each group of 8PSK modulation for the one-receive-antenna case. Simulation results exactly matched our ST-BICM performance analysis. Also, these results confirmed the validity of the proposed design guidelines, and provided intuition for the performance evaluation. From the analysis, we have shown that the small gain groups and the large gain groups cross over in the BER curves. Therefore, we can use the proposed design guideline such that a proper mapping can be selected depending on the system requirements.

Extension to ST block-coded systems is studied in [12]. The mapping optimization for higher modulation levels, such as 16QAM, remains as a future research topic.

Appendix

Proof of Equation (2): The bit-error probability for punctured convolutional codes is bounded by [9]

$$P_b \le \frac{1}{T_p} \sum_{d=d_H}^{\infty} c_d f(d,\mu,\chi) \tag{4}$$

where T_p denotes the puncturing period, c_d represents the total input weight of error events at Hamming distance d, and $f(d,\mu,\chi)$ indicates the pairwise error probability (PEP) of coded systems.

It has been shown in [1] and [2] that the performance of iterative decoding schemes approaches maximum-likelihood (ML) performance with the assumption of exact feedback from the decoder. Following the analysis in [8], the PEP is obtained by

$$f(d,\mu,\chi) \le (mN_t)^{-d} \sum_{\underline{\mathbf{S}}} 2^{-d} \sum_{\underline{\mathbf{U}}} 2^{-d(m-1)} \sum_{\mathbf{x}\in\chi_{\underline{c}}^{\underline{\mathbf{S}}}} P(\mathbf{x}\to\mathbf{z}) \tag{5}$$

where $\mathbf{x}$, $\mathbf{z}$ denote the transmitted and erroneously detected sequences, respectively, $\underline{\mathbf{S}}$ represents a random process whose elements' outcome determines the switch position (identically independent distributed, uniformly distributed over $\{(1,1),\ldots,(i,n),\ldots,(N_t,m)\}$ with the antenna index i and the bit index n), $\underline{\mathbf{U}}$ denotes a binary random process whose element for each transmission over the vector channel selects either μ or $\bar{\mu}$, independently with probability 1/2, with $\bar{\mu}$ defined as the mapping obtained by complementing μ, and $\chi_{\underline{c}}^{\underline{\mathbf{S}}}$ indicates the sequence of signal subsets selected by the correct bits [8].

Denote $N = N_v/m$ as the symbol length of the transmitted sequence. Assuming that $\mathbf{x}$ is transmitted, we have the average SEP that ML chooses $\mathbf{z}$ over $\mathbf{x}$ as

$$P(\mathbf{x}\to\mathbf{z}|\mathbf{H}) = P\left(\sum_{t=1}^{N}\left(\|\mathbf{y}_t-\mathbf{H}_t\mathbf{x}_t\|^2-\|\mathbf{y}_t-\mathbf{H}_t\mathbf{z}_t\|^2\right)>0\right).$$

Using a Chernoff bound [13], this average SEP is bounded by

$$\begin{aligned} &P(\mathbf{x}\to\mathbf{z}|\mathbf{H}) \\ &\le \min_{\lambda>0}\exp\left\{\lambda\left(\sum_{t=1}^{N}\left(\|\mathbf{y}_t-\mathbf{H}_t\mathbf{x}_t\|^2-\|\mathbf{y}_t-\mathbf{H}_t\mathbf{z}_t\|^2\right)\right)\right\} \\ &= \min_{\lambda>0}\prod_{k=1}^{d}\prod_{j=1}^{N_r}\exp\left\{\lambda\left(\left|y_k^j-h_k^{c_k,j}x_k^{c_k}\right|^2\right.\right. \\ &\qquad\left.\left.-\left|y_k^j-h_k^{c_k,j}z_k^{c_k}\right|^2\right)\right\} \end{aligned}$$

where k and c_k denote the time index and the transmitter index of the erroneously transmitted symbol, respectively, and λ indicates the Chernoff bound parameter. In the above derivation, we consider only one symbol error out of each erroneously transmitted vector with the exact feedback assumption from the decoder.

Minimizing over λ and averaging with respect to independent Rayleigh distributions of $h_k^{c_k,j}$, the average SEP is expressed as

$$P(\mathbf{x}\to\mathbf{z}) \le \prod_{k=1}^{d}\prod_{j=1}^{N_r}\frac{1}{1+|x_k^{c_k}-z_k^{c_k}|^2/2\sigma^2}.$$

Here, we assume the same signal subset χ for all transmit antennas at each transmission. Thus, we will omit the indices k and c_k.

Applying these results to (5), we obtain

$$f(d,\mu,\chi) \le \left(\frac{1}{m\cdot 2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{x\in\chi_b^i}\left(\frac{1}{1+|x-z|^2/2\sigma^2}\right)^{N_r}\right)^d$$

where χ_b^i is denoted by $\{\mu([v^1,v^2,\ldots,v^m])|v^i=b\}$ for any transmit antenna.

Finally, we substitute the above equation into the bound (4) and apply a logarithmic scale on both sides to result in (2).

REFERENCES

[1] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1250–1257, Aug. 2002.

[2] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. Veh. Technol. Conf.*, Sep. 2000, vol. 1, pp. 473–478.

[3] Y. Huang and J. A. Ritcey, "Improved 16-QAM constellation labeling for BI-STCM-ID with the Alamouti scheme," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 157–159, Feb. 2005.

[4] F. Schreckenbach and G. Bauch, "Irregular signal constellations, mappings and precoder," in *Proc. Int. Symp. Inf. Theory Appl.*, Oct. 2004, CD-ROM.

[5] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[6] I. Lee, A. M. Chan, and C.-E. W. Sundberg, "Space-time bit-interleaved coded modulation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 820–825, Mar. 2004.

[7] I. Lee and C.-E. W. Sundberg, "Wireless OFDM systems with multiple transmit and receive antennas with bit interleaved coded modulation," *IEEE Wireless Commun.*, vol. 13, pp. 80–87, Jun. 2006.

[8] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 927–946, May 1998.

[9] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389–400, Apr. 1988.

[10] S. Benedetto, G. Montorsi, and D. Divsalar, "Concatenated convolutional codes with interleavers," *IEEE Commun. Mag.*, no. 8, pp. 102–109, Aug. 2003.

[11] J. M. Cioffi, Lecture Notes for EE379A: Signal Processing and Detection Stanford Univ.. Stanford, CA, Jan. 2005 [Online]. Available: http://www.stanford.edu/class/ee379a

[12] J. Choi, W. Lee, and I. Lee, "Mapping optimization for space-time block coded OFDM systems with iterative decoding," in *Proc. Veh. Technol. Conf.*, May 2005, vol. 2, pp. 1230–1234.

[13] D. Divsalar and M. K. Simon, "Trellis coded modulation for 4800-9600 bits/s transmission over a fading mobile satellite channel," *IEEE J. Sel. Areas Commun.*, vol. SAC-5, no. 2, pp. 162–175, Feb. 1987.

[14] I. Lee and C.-E. Sundberg, "Reduced-complexity receiver structures for space–time bit-interleaved coded modulation systems," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 142–150, Jan. 2007.