Coded Modulation with Signal Space Diversity

Qiuliang Xie, Jian Song, Kewu Peng, Fang Yang, and Zhaocheng Wang

***Abstract*—Signal space diversity (SSD) is a well-known power- and bandwidth-efficient diversity technique with constellation rotation and coordinate interleaving. This paper investigates issues of combining SSD with coded modulation systems (CMSs), e.g., bit-interleaved coded modulation (BICM) and its iterative version, BICM-ID. It demonstrates that the average mutual information (AMI) between the signal after rotated constellation mapping and the signal before or after the soft demapper varies with the rotation angle. A new criterion for determining the optimal rotation angle by maximizing such AMI is therefore proposed. Specific considerations of combining SSD with BICM (BICM-SSD) and BICM-ID (BICM-ID-SSD) are addressed. The demapper's extrinsic information transfer (EXIT) curve in traditional BICM-ID systems exhibits different slopes under different channels, which consequently prevents traditional BICM-ID systems from having simultaneously excellent performance under different channels. It is shown that such a different-slope problem can be simply mitigated by a proper use of SSD. Analysis and simulation show that the proposed BICM-ID-SSD systems hold a near-capacity performance under both additive white Gaussian noise (AWGN) and Rayleigh fading channels simultaneously.**

***Index Terms*—Signal space diversity, average mutual information, BICM, BICM-ID, EXIT chart.**

I. Introduction

ORIGINALLY proposed by J. Boutros and E. Viterbo [1], signal space diversity (SSD) is a promising technique providing significant diversity gain over fading channels without any bandwidth penalty, meanwhile without any degradation over additive white Gaussian noise (AWGN) channels. The basic idea of SSD is to use coordinate interleaving (also named the in-phase and quadrature phase (I/Q) interleaving for two dimensional constellations), together with constellation rotation to increase the diversity order over fading channels.

The determination of the optimal constellation rotation angle for SSD is an open problem. Many papers tried to solve this problem, which can be classified mainly into two categories: rotation for uncoded systems [1], [2], [3], [4], [5] and for coded systems [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. As practical systems always require channel coding, optimization work for coded system is more important. The most promising coded modulation systems (CMSs) over fading channels are bit-interleaved coded modulation (BICM) [16] and its iterative version, BICM with iterative demapping and decoding (BICM-ID) [17], [18].

Manuscript received June 1, 2010; revised September 2, 2010; accepted November 3, 2010. The associate editor coordinating the review of this paper and approving it for publication was A. Sezgin.

Part of this paper was presented at IEEE Globecom, 2010. This work was supported by the Chinese AQSIQ Project 200910244.

The authors are with the Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Electronic Engineering, Tsinghua University, Beijing 100084, P. R. China (e-mail: xql06@mails.tsinghua.edu.cn, xieqiuliang@gmail.com).

Digital Object Identifier 10.1109/TWC.2011.120810.100951

Combining BICM with SSD (BICM-SSD) is discussed in [6], [7], [8]. Chindapol and Ritcey [6] tried to find the optimal rotation angles for BICM-SSD systems over Rayleigh fading channels, wherein 1/2-rate convolutional codes with 4/16 quadratures amplitude modulation (QAM) and 8 phase shift keying (PSK) were considered. BICM-SSD with limit-approaching low-density parity-check (LDPC) codes was presented by Kiyani *et.al* [7] with PSK over Rayleigh channels. Nour and Douillard [8] investigated LDPC coded BICM-SSD, wherein Rayleigh channels with/without erasures and the 4/5 and 9/10-rate LDPC codes of DVB-T2 were considered, as LDPC coded BICM-SSD is one of the key technologies of DVB-T2 [19], [20]. However, optimization criteria in the above literatures are based on the *harmonic mean* of the square of Euclidean distance [6] similar to that for uncoded systems [1], or based on extensive computer simulation [7], [8], which are not effective enough indeed. Moreover, it is interesting that for Shannon-limit approaching BICM-SSD systems over fading channels, no attention has been paid to high order QAM at low or moderate code rates, which will be explained later in this paper.

Combining BICM-ID with SSD (BICM-ID-SSD) is extensively addressed in [9], [10], [11], [12], [13], [14], [15]. Chindapol and Ritcey [9] considered BICM-ID-SSD with square QAM but without rotation over Rayleigh fading channels. It showed that significant diversity order can be achieved by SSD even without constellation rotation. Kiyani and Weber [11], [12] mainly tried to find the optimal rotation angle via exhaustive computer simulation, and claimed that the optimal rotation angle is dependent upon labeling functions. Tran *et.al.* used the bit error rate (BER) bound as the criterion to determine the optimal rotation matrices [13], [14], [15]. Nevertheless, computer simulation costs too much time yet the results can not be generalized, as they depend on too many specific factors such as the channel coding and labeling. The BER bound is a good criterion, but it also depends on these factors, and thus makes the optimization problem complicated. Moreover, the BER bound is not suitable for noisy channels at low or moderate signal to noise ratio (SNR) because it is a good approximation only at high SNR, and therefore can not provide a useful guideline for near-capacity CMS design.

In this paper, it shows that the average mutual information (AMI) between the signal after rotated constellation mapping and the signal before/after the soft demapper varies with the rotation angle over fading channels, in the case of perfect channel state information available at the receiver side (CSIR). Therefore, AMI-maximization is proposed as the criterion in searching for the optimal rotation angle. The basic idea of this paper is summarized as follows. Since the AMI bounds the maximum information rate that can be transmitted reliably at a given SNR, or the lowest SNR limit above which a given

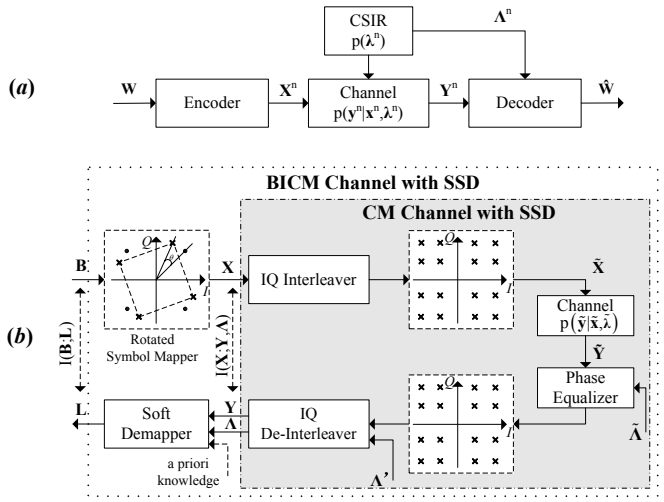

Fig. 1. System models: (a) a general model for point-to-point communication systems with CSIR, and (b) a model of typical communication systems with SSD.

rate of information that can be reliably transmitted, we first try to maximize the AMI at a given SNR, or equivalent, to minimize the SNR limit at a given AMI. After that, we try to design a practical CMS to approach such SNR limit following the iterative principle.

The rest of this paper is organized as follows. Section II deals with the rotation problem for both optimal CMSs (successive or iterative demapping systems, e.g., BICM-ID) and independent demapping CMSs (e.g., BICM), based on the AMI criterion using a general SSD system model. The overall optimization framework is then proposed. Section III tries to approach the optimized limits derived in Section II. For BICM-ID-SSD systems, the extrinsic information transfer (EXIT) chart [21] is employed to help design, where the doping technique is also employed for error-floor removal [22]. Simulation is carried out in Section IV to demonstrate the high efficiency of the optimization framework. Conclusions are drawn in Section V.

Throughout this paper we use capital letters to represent random variables and the corresponding small letters to represent their realizations. Bold letters denote vectors or matrices. $\text{SNR} = E_s/N_0$ where E_s denotes the symbol energy and N_0 denotes the variance of the complex Gaussian noise. Code rates less than $3/4$ (especially, around $1/2$) are of interests in this paper, because the performance of low-order constellation with high code rates is much worse than that of high-order constellation with low to moderate code rates to achieve the same spectrum efficiency, which will be detailed later via AMI analysis.

II. System Models and AMI Analysis

A. Basic System models

A general model of point-to-point communication systems with CSIR is shown in Fig. 1 (a), where $(\cdot)^n$ denotes a vector with length n. The Shannon capacity (or ergodic capacity) can be written as [23]

$$C = \max_{p(\boldsymbol{x})} I(\boldsymbol{X}; \boldsymbol{Y}|\boldsymbol{\Lambda}). \tag{1}$$

for a memoryless channel where $I(\cdot;\cdot)$ represents the AMI function. Throughout this paper we refer to memoryless channels, typically for example, the identically independently distributed (i.i.d.) Rayleigh fading channels. The system model shown in Fig. 1 (a) is basically of interests from the information theory point of view, and the channel capacity is the maximum AMI between $\boldsymbol{X}$ and $\boldsymbol{Y}$ with the CSIR $\boldsymbol{\Lambda}$ by adjusting the distribution of $\boldsymbol{X}$, under some cost constraints such as the power constraint for channels with continuous-alphabet (typically, the AWGN channel). However, for a practical communication system, the transmitted signal $\boldsymbol{X}$ is additionally constrained by other constraints such as usually $\boldsymbol{X}$ takes on a discrete finite alphabet χ (the constellation signal set) with equiprobability.

The system model of SSD is depicted in Fig. 1 (b). As shown in it, an m-tuple encoded bit vector $\boldsymbol{B} = [B_0 B_1 \cdots B_{m-1}]^t$ is mapped on to a constellation symbol $\boldsymbol{X}$, where the constellation may be rotated to achieve better performance, and the superscript ^{t} denotes transposition. After that, the I/Q components of the mapped symbols are interleaved. New symbols are then reconstructed based on these interleaved I/Q components and transmitted over the channel. With perfect CSIR, the received signal is first phase-equalized and then I/Q de-interleaved before sending to the soft demapper. Assuming that the interleaver is long and random enough, which can be implemented via the cyclic Q-delay and symbol interleaving [9], each coordinate of the symbol sent to the demapper can be regarded as suffering from independent fading and the equivalent coded modulation (CM) channel with SSD shown in the dot-dashed shadow box in Fig. 1 (b) can be modeled as

$$Y_K = \Lambda_K X_K + N_K, K \in \{I, Q\} \tag{2}$$

where Λ_I and Λ_Q are i.i.d. fading coefficients. N_I and N_Q are i.i.d. Gaussian random noise with zero mean and variance of $N_0/2$. X_I and X_Q denote the I/Q components of the rotated symbol, respectively. For Rayleigh fading channels, Λ_I and Λ_Q are Rayleigh distributed with the normalized probability density function (PDF) of $p(\lambda) = 2\lambda \exp(-\lambda^2), \lambda \geq 0$. By denoting $\boldsymbol{X} = [X_I, X_Q]^t$, $\boldsymbol{Y} = [Y_I, Y_Q]^t$, $\boldsymbol{N} = [N_I, N_Q]^t$ and $\boldsymbol{\Lambda} = \text{diag}(\Lambda_I, \Lambda_Q)$ representing a (2×2) diagonal matrix with the diagonal elements Λ_I and Λ_Q, the channel modeled in (2) can also be written in the matrix form as $\boldsymbol{Y} = \boldsymbol{\Lambda X} + \boldsymbol{N}$.

B. AMI Analysis

As the coded modulation (CM) channel with SSD can be modeled by (2), the constellation constraint AMI with perfect CSIR over a particular fading channel, called the CM-AMI since it bounds the maximal information rate that can be reliably transmitted for all CMSs with equiprobable inputs, could be written and evaluated as [24]

$$\begin{aligned} I_{\text{CM}} &= I(\boldsymbol{X}; \boldsymbol{Y}|\boldsymbol{\Lambda}) \\ &= m - \mathbb{E}_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda}} \left[\log_2 \frac{\sum_{\hat{\boldsymbol{x}} \in \chi} p(\boldsymbol{y}|\hat{\boldsymbol{x}}, \boldsymbol{\lambda})}{p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\lambda})} \right], \end{aligned} \tag{3}$$

where χ denotes the rotated constellation set, $m = \log_2 |\chi|$ whereby $|\chi|$ denotes the size of the set and $\mathbb{E}[\cdot]$ denotes expectation. Since the channel input $\boldsymbol{x}$ in (3) is constrained

by set χ, different rotation angle results in different set χ and consequently different CM-AMI.

The AMI after independent demapping which is relevant to the labeling, called BICM-AMI, is defined as [24]

$$I_{\text{BICM}} = \sum_{k=0}^{m-1} I(B_k; \boldsymbol{Y}|\boldsymbol{\Lambda}). \tag{4}$$

Lemma 1: $I(B_k; \boldsymbol{Y}|\boldsymbol{\Lambda}) = I(B_k; L_k), \forall k \in \{0, \cdots, m-1\}$ after the maximum *a posteriori* (MAP) demapping, i.e.,

$$L_k = \log \frac{\Pr(B_k = 0|\boldsymbol{Y}, \boldsymbol{\Lambda})}{\Pr(B_k = 1|\boldsymbol{Y}, \boldsymbol{\Lambda})}. \tag{5}$$

Proof: As $I(B_k; \boldsymbol{Y}|\boldsymbol{\Lambda}) = H(B_k) - H(B_k|\boldsymbol{Y}, \boldsymbol{\Lambda})$ and $I(B_k; L_k) = H(B_k) - H(B_k|L_k)$ where $H(\cdot)$ denotes the entropy function, it is equivalent to show $H(B_k|\boldsymbol{Y}, \boldsymbol{\Lambda}) = H(B_k|L_k)$. In fact, $\Pr(B_k = b|L_k) = \Pr(B_k = b|\boldsymbol{Y}, \boldsymbol{\Lambda}) = \exp[L_k(1-b)]/[1+\exp(L_k)], \forall b \in \{0, 1\}$ according to the MAP demapping method (5). Therefore, it is apparently that $H(B_k|\boldsymbol{Y}, \boldsymbol{\Lambda}) = H(B_k|L_k)$ and Lemma 1 is proved.

Apparently labeling has a significant influence on BICM-AMI and Gray labeling maximizes it [24]. However, based on the proof shown above, Lemma 1 holds true for any labeling function. This lemma gives an effective way to calculate the BICM-AMI by measuring the AMI between the transmitted bits and their corresponding LLRs, i.e.,

$$\begin{aligned} I_{\text{BICM}} &= \sum_{k=0}^{m-1} I(B_k; L_k) \\ &= m + \sum_{k=0}^{m-1} \mathbb{E}_{L_k}\left[f_I(L_k)\right], \end{aligned} \tag{6}$$

where $f_I(L_k) = \frac{e^{L_k}}{1+e^{L_k}} \log_2 \frac{e^{L_k}}{1+e^{L_k}} + \frac{1}{1+e^{L_k}} \log_2 \frac{1}{1+e^{L_k}}$. Also, different rotation angle results in different BICM-AMI, because the output signal $\boldsymbol{Y}$ in (4) is relevant to the input signal $\boldsymbol{X}$ which is constrained by the rotated signal set χ.

C. Important Properties and the Optimization Framework

1) Labeling: Typical labeling functions are one-to-one functions. Therefore, $I(\boldsymbol{B}; \boldsymbol{Y}|\boldsymbol{\Lambda}) = I(\boldsymbol{X}; \boldsymbol{Y}|\boldsymbol{\Lambda})$. Since $I(\boldsymbol{X}; \boldsymbol{Y}|\boldsymbol{\Lambda})$ is independent with labeling functions, it implies that the CM-AMI is not relevant to labeling even from the bit-level viewpoint. Therefore, the optimal constellation rotation angle which maximizes the CM-AMI is independent with the labeling for optimal CMSs, e.g., BICM-ID-SSD systems. However, for BICM-SSD systems, the BICM-AMI given by (4) strongly depends on the labeling. Numerous labeling functions exist for high-order constellations. Fortunately, constellations with Gray labeling are optimal for BICM systems over AWGN channel [25], [24]. Thereby, we mainly focus on Gray labeling because the performance over AWGN channel should not be degraded, even sometimes Gray labeling may not be the best for BICM-SSD systems over fading channels, which will be detailed later.

2) Square QAM without Rotation: For square QAM without rotation, X_I and X_Q are independent with each other, i.e., $\Pr\{X_I = x_I, X_Q = x_Q\} = \Pr\{X_I = x_I\}\Pr\{X_Q = x_Q\}$. Therefore, the I/Q channels are independent with each other that $I(\boldsymbol{X}; \boldsymbol{Y}|\boldsymbol{\Lambda}) = I(X_I; Y_I|\Lambda_I) + I(X_Q; Y_Q|\Lambda_Q)$, which implies that the CM-AMI will be unchanged whether I/Q interleaving is employed or not. For independent demapping systems with Gray labeling, since all of its binary signal sets lie vertically or horizontally, the BICM-AMI with I/Q interleaving also remains exactly the same. Therefore, I/Q interleaving does not change the AMI for square QAM without rotation, for both BICM-ID-SSD and Gray labeled BICM-SSD systems.

3) Without IQ Interleaving or over AWGN Channels: We have $\Lambda_I \equiv \Lambda_Q$ without I/Q interleaving and $\Lambda_I \equiv \Lambda_Q \equiv 1$ over AWGN channel. Therefore, the AWGN case can be regarded as a special case of those without I/Q interleaving, under which condition the received signal $\boldsymbol{Y}$ can be equivalently rotated back to the traditional signal, and the noise distribution remains the same. As a result, both CM-AMI and BICM-AMI will be unchanged if only constellation rotation is employed or over AWGN channel.

4) Link with EXIT Charts: There are several properties of EXIT charts [26]. One important property is that the area under the demapper's EXIT curve $\mathcal{A}_{\text{dem}}$ approximately equals to the CM-AMI I_{CM} over the number of bits per constellation point m [26], [27], shown as

$$\mathcal{A}_{\text{dem}} \approx I_{\text{CM}}/m. \tag{7}$$

Based on this property, once we have a larger CM-AMI via SSD, we have a larger area, which imply that it would be easier to let the demapper's EXIT curve to be above the outer decoder's inverted EXIT curve for BICM-ID-SSD system design.

Another property is that, without *a priori* information input, the extrinsic information output from the MAP demapper IE_0 shown in (5) exactly equals to the BICM-AMI I_{BICM} over m based on Lemma 1, i.e.,

$$\text{IE}_0 = I_{\text{BICM}}/m. \tag{8}$$

5) Optimization Framework: Since the CM-AMI is not relevant to labeling or outer channel codes, and only Gray labeling deserves analysis for BICM, the CMS design optimization can be broken into two independent steps. First, try to maximize the AMI, either CM-AMI for BICM-ID or BICM-AMI for BICM. This AMI maximization allows the maximum information rate that can be reliably transmitted at a given SNR, or the lowest SNR limit above which a given rate of information that can be reliably transmitted. Second, try to approach the limit via turbo (iterative) principle [28]. For BICM systems, this task turns to be a code design problem and has been discussed in many papers, so that no more efforts will be paid here. For BICM-ID systems, details will be given in the next section using EXIT charts.

D. Numeric Results

Numeric results are provided in searching for the optimal rotation angle over Rayleigh channels. Both I_{CM} and Gray

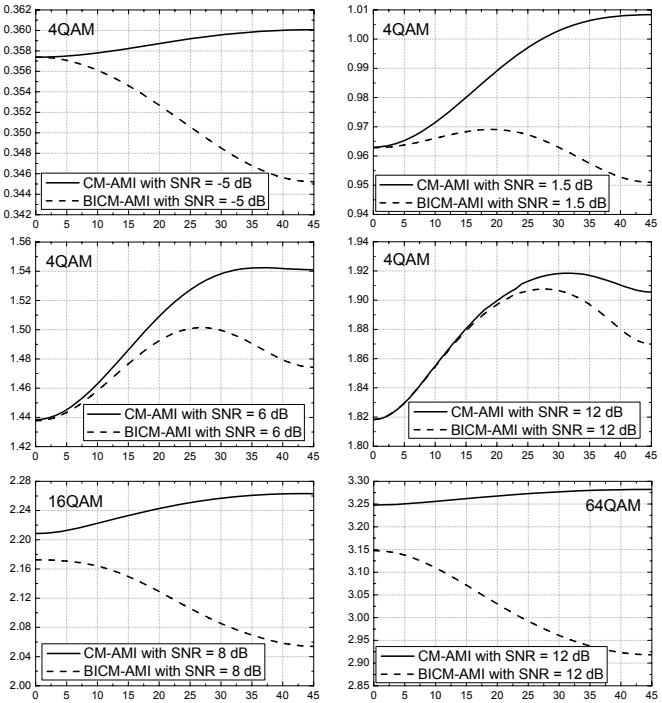

Fig. 2. CM-AMI I_{CM} and BICM-AMI I_{BICM} vs. rotation angle θ for 4/16/64QAM at different SNR over Rayleigh fading channels. The rotation angle in degree is chosen as the X-axis, and the AMI in bits/channel use is chosen as the Y-axis.

labeled I_{BICM} associated with 4/16/64QAM are presented. Given a constellation and a fading channel, I_{CM} is a function of the rotation angle θ and the SNR, i.e., $I_{\text{CM}}(\theta, \text{SNR})$, and I_{BICM} is also a function of θ and SNR with a fixed labeling, i.e., $I_{\text{BICM}}(\theta, \text{SNR})$.

We first fix the SNR and plot how $I_{\text{CM}}(\theta, \text{SNR})$ and $I_{\text{BICM}}(\theta, \text{SNR})$ vary with the rotation angle θ. The numeric results are shown in Fig. 2. As shown in this figure for 4QAM, the rotation angle of $\theta = 45^\circ$ can be regarded as optimal at low to moderate SNR, e.g., at SNR $= 1.5$ dB or SNR $= 6$ dB, the CM-AMIs are about 1 bit/channel use or 1.5 bits/channel use corresponding to the 1/2 and 3/4 code rates, respectively. However, the optimal rotation angle for BICM-AMI varies a lot with SNR or code rates. Taking 4QAM as an example, at SNR $= -5$ dB, $\theta = 0$ is optimal, while at SNR $= 1.5$ dB, the optimal θ is about 20°, and at SNR $= 6$ or 12 dB, such optimal θ is about 27°. The situations for higher order constellations are similar. Another interesting phenomenal can be observed that, 45° is also optimal for both 16/64QAM at code rates of practical interests from the CM-AMI point of view, while from the BICM-AMI viewpoint, no rotation is optimal for both 16/64QAM. This observation explains why there are no papers discussing the rotation problem with high order constellations at low to moderate code rates for BICM-SSD systems, with limit-approaching codes over Rayleigh channels.

We now fix the rotation angle and plot how $I_{\text{CM}}(\theta, \text{SNR})$ and $I_{\text{BICM}}(\theta, \text{SNR})$ vary with SNR. For CM-AMI, the rotation angle is fixed as 45°. For BICM-AMI, $\theta = 22.5^\circ$ is chosen for 4QAM and $\theta = 0$ for 16/64QAM. Since the AMI values are too close to each other, the AMI is chosen as the X-axis, and the gap to the Gaussian input as the Y-axis, via which distribution the channel capacity with CSIR can be achieved and evaluated as $C = \mathbb{E}_\lambda \log_2(1 + \lambda^2 \text{SNR})$. The numeric results over Rayleigh fading channels are plotted in Fig. 3, based on which several interesting observation can be made as follows. First, the CM-AMI can be increased via rotation, especially for low order constellations. Second, independent demapping loss is considerable at low to moderate code rates for high order constellations. Therefore, the iterative or successive demapping technique is highly recommend to be used when SSD is employed. Third, from the CM-AMI point of view, higher order constellations are always better than the lower order constellations. From the BICM-AMI point of view, this figure provides a reference to help choose the system parameters. For instance, at the AMI of 1.5 bit/channel use, it is highly recommended to choose 16QAM combined with a rate-3/8 code, rather than 4QAM with a rate-3/4 code or 64QAM with a rate-1/4 code. On the other hand, the scheme of using higher order constellations together with low code rates relatively results in high complexity. Therefore, we would like to state that moderate code rates around 1/2 are of practical interests.

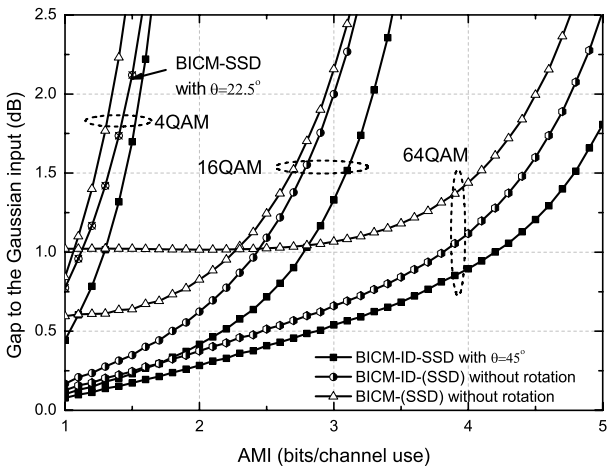

Fig. 3. Gap to the Gaussian input channel capacity over Rayleigh fading channels.

As shown in Fig. 2, although the CM-AMI with constellation rotation is larger than that without rotation for all 4/16/64QAM, the BICM-AMI with rotation is worse for 16/64QAM at practically interested code rates and for 4QAM at low code rates. This phenomena implies that the independent demapping loss in BICM-SSD systems with constellation rotation is too large at low to moderate code rates. Gray labeling is known optimal for BICM without SSD. However, for BICM-SSD system with constellation rotation, it may not be optimal over fading channels. Taking square QAM associated with the rotation angle of $\theta = 45^\circ$ as an example, the received signal $\boldsymbol{Y}$ is no more square. The distance between the diagonal points in square QAM may sometimes be smaller than that between the traditional nearest neighbors, with the probability of $P = 1 - \Pr\left\{\sqrt{2}/2 \le \Lambda_I/\Lambda_Q \le \sqrt{2}\right\} = 2/3$ over Rayleigh fading channels.

III. EXIT Chart Analysis

A. SSD Effect to the Demapper's EXIT Curve

EXIT chart is a powerful tool for analyzing the convergence behavior of iterative systems [21]. It depicts the extrinsic

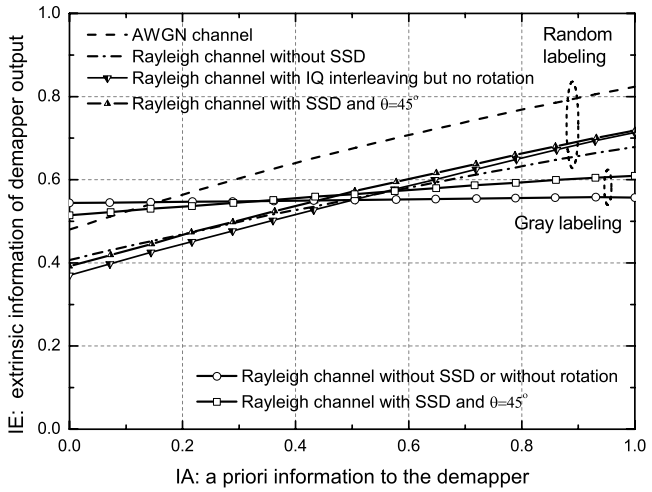

Fig. 4. SSD effect to the 16QAM demapper's EXIT curve at SNR = 8 dB.

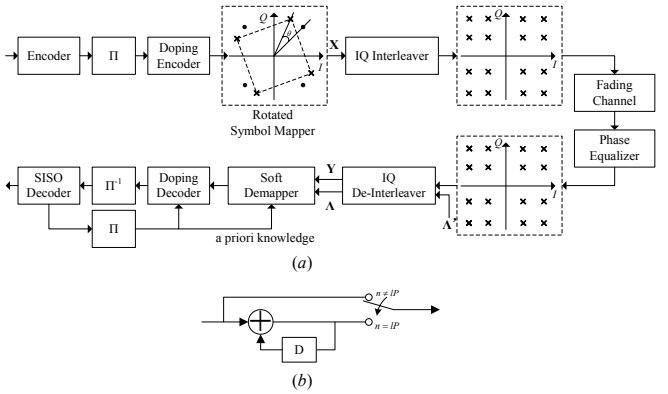

Fig. 5. (a) BICM-ID-SSD system model with doping. (b) The inner RSC code with doping: every Pth information bit is replaced by a coded redundant bit.

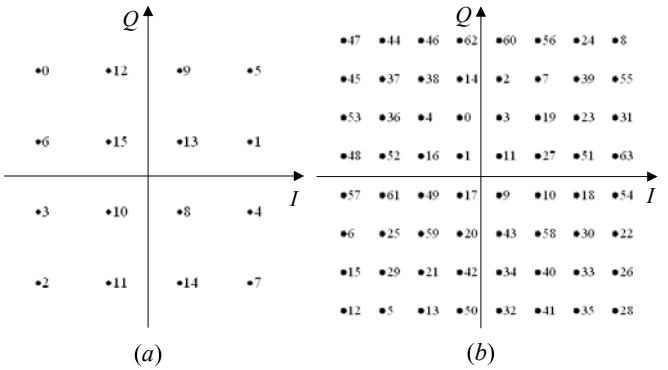

Fig. 6. Constellation labeling. (a) 16QAM and (b) 64QAM.

information transferred between iterative modules and has found a wide applications, due to its high accurate prediction and low computation cost. There are many properties of EXIT charts [26]. The area property shown in (7) and the start point property of the demapper's EXIT curve shown in (8) are demonstrated in Fig. 4, using 16QAM over Rayleigh fading channels as an example. It can be observed in Fig. 4 that the area under the demapper's EXIT curve with constellation rotation is larger than that without rotation. This observation matches well with the result shown in Fig. 2 that the CM-AMI can be increased by constellation rotation. Another observation is that, with Gray labeling at the SNR of 8 dB, the start point (without *a priori* information) of the demapper's EXIT curve with rotation is lower than that without rotation, as shown in Fig. 4, which also matches well with the result shown in Fig. 2 that constellation rotation may results in negative influence on the BICM-AMI for high-order constellations. Furthermore, a more important observation can be made that SSD may let the demapper's EXIT curve to have a larger slope, which will be addressed in the following.

It is well-known that EXIT curve-fitting is an essential way for BICM-ID system design, i.e., to choose a pair of well-matched labeling and outer channel code from the EXIT chart point of view. However, the demapper's EXIT curve has different slopes under different channels. For instance, the demapper's EXIT curve under Rayleigh fading channels (the dot-dashed line) has a smaller slope than that under AWGN channel (the dashed line) without SSD, as shown in Fig. 4 taking 16QAM as an example. Therefore, a well-matched pair of labeling and outer code under AWGN channel may not match well under fading channels simultaneously. This different-slope-problem was also observed in other papers [29], but no general methods have been proposed to solve it. Fortunately, we show that SSD can change the demapper's EXIT curve to have a larger slope under fading channels, even without constellation rotation for most labeling functions [30]. Labeling with all its binary signal sets lying vertically or horizontally, such as the Gray labeled square QAM without rotation, does not benefit from this property. The following lemma shows that SSD plays a very important role for a BICM-ID system to exhibit simultaneously excellent performance under different channels.

Lemma 2: SSD effect to the demapper's EXIT curve.

1) With the same CM-AMI, the demapper's EXIT curve under fading channels without SSD has a smaller slope than that under AWGN channels.
2) SSD can change the demapper's EXIT curve to have a larger slope or let the curve to be above the original curve without SSD.
3) To let the demapper's EXIT curve under fading channels exhibit approximately the same slope as that under AWGN channel, it is no need to use very high-dimensional constellations combining with SSD.

Proof: see in the Appendix.

B. BICM-ID-SSD System Design Based on EXIT Charts

For BICM-SSD system design, after determining the optimal rotation angle by maximizing the BICM-AMI, the remaining work is just to choose a powerful code such as a turbo or LDPC code to achieve excellent performance. However, for BICM-ID-SSD system design, finding a pair of well-matched labeling and outer channel code remains a big challenge, after the rotation angle optimization according to the CM-AMI maximization.

The doping technique is employed here for error-floor removal, which is originally proposed in [22] for BICM-ID systems. The system model of BICM-ID-SSD with doping

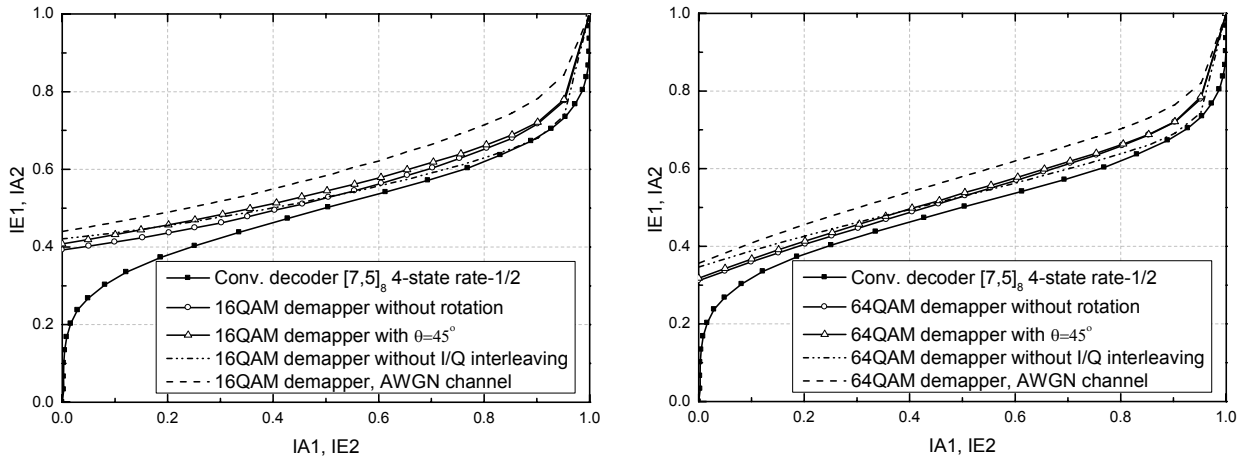

Fig. 7. EXIT chart analysis for BICM-ID-SSD systems with 16/64QAM labeling shown in Fig. 6. The doping rate is chosen as 100. (a) For 16QAM, SNR $=$ 8 dB over Rayleigh fading channel and SNR $=$ 7 dB over AWGN channel. (b) For 64QAM, SNR $=$ 12 dB over Rayleigh fading channel and SNR $=$ 11 dB over AWGN channel.

is shown in Fig. 5 (a) by adding a doping code before the constellation symbol mapper. The doping code we use is a unity-rate 2-state recursive systematic convolutional (RSC) code, whose encoder is depicted in Fig. 5 (b) with every Pth information bit being replaced by a coded redundant bit, where P is called the *doping rate*, which is chosen as 100 in this paper.

The outer channel code we choose is a 1/2-rate 4-state non-recursive convolutional (NRC) code with the generator of $[7, 5]_8$ where the compact notation NRC$([G_1, G_2]_8)$ represents a NRC code with feed forward polynomials G_1 and G_2 in octal. Labeling is well-known crucial for BICM-ID system design [9], [17], [13], [14], [15], [31], [32], [33], [34], [35], [36], [37]. Based on the binary switch algorithm (BSA) [36], [37], labeling for 16/64QAM is found and depicted in Fig. 6.

The EXIT charts are provided in Fig. 7. The outer convolutional code uses the standard Bahl-Cocke-Jelinek-Raviv (BCJR) decoding algorithm, and the so-called *serial detection* method [22] is used for doping decoding. As shown in this figure, the doped demapper's EXIT curve matches the outer decoder's inverted EXIT curve very well under both AWGN and Rayleigh fading channels when SSD is employed. Therefore, BICM-ID-SSD systems with excellent performance can be expected under both channels. It is worth noting that the error floor in traditional BICM-ID systems can be removed by the doping technique, and the different-slope-problem of the demapper's EXIT curve under different channels can be mitigated by using SSD.

IV. Simulation Results

A. Results of BICM-SSD Systems

The parameters of the simulated BICM-SSD system with 4QAM is described as follows. The rotation angle is chosen as 22.5° for both 1/2 and 3/4 code rates, and the DVB-T2 64,800-bit-long LDPC codes [19] are selected. The MAP demapping is employed, and the LDPC decoders use the standard *sum-product algorithm* [38] with a maximal iteration of 100 times. The simulation results are presented in Fig. 8. As shown in this figure, about 0.2 dB can be achieved via SSD at the code rate of 1/2, and about 0.8 dB at the rate of 3/4, both at BER of 10^{-5}. These results verify the efficiency of the BICM-AMI analysis shown in Fig. 3.

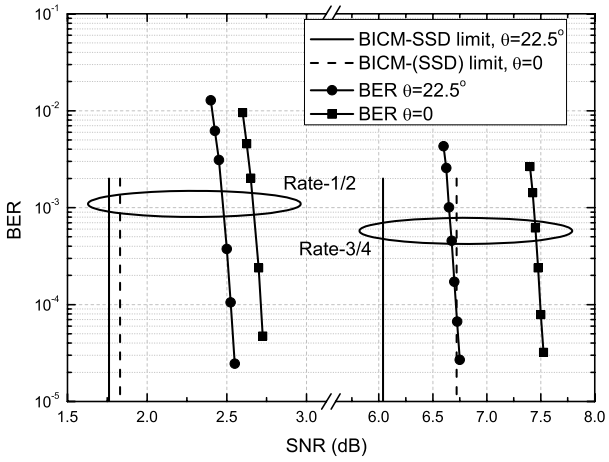

Fig. 8. BER simulation results of BICM-SSD systems with 4QAM over Rayleigh fading channels. DVB-T2 LDPC codes with code length of 64,800 bits are used.

For 16/64QAM, our simulation results (although not presented here) show that BICM-SSD with rotation performs worse than that without rotation, at the code rates of practical interests, which also verifies the efficiency of the BICM-AMI analysis.

B. Results of BICM-ID-SSD Systems

The block length of the proposed BICM-ID-SSD systems is set as 64,800 bits. Reference BICM systems with the 1/2-rate 64,800-bit-long DVB-T2 LDPC code are also simulated. The length of the bit-interleaver employed in these BICM systems is 324,000 bits, i.e., 5 LDPC-codeword long, while the bit-interleaver in the BICM-ID-SSD systems is an S-random interleaver with the length of 64,800 bits [39], wherein the parameter S is chosen as 100, equaling to the doping

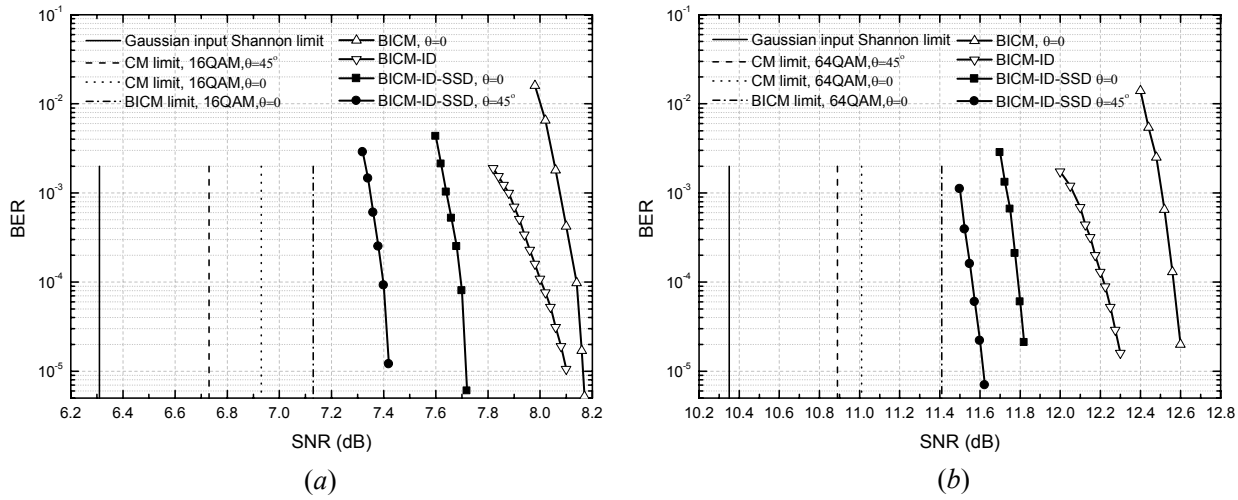

Fig. 9. BER performance of the proposed BICM-ID-SSD system with/without constellation rotation, at the code rate of $1/2$ over Rayleigh fading channels. The performance of BICM system with the 1/2-rate DVB-T2 LDPC code is also provided as reference. (a) 16QAM and (b) 64QAM.

rate P. Both the proposed BICM-ID-SSD system and the LDPC code take a maximal iteration of 100 times. The I/Q interleaver is assumed to be long and random enough making each coordinate suffering from independent Rayleigh fading. Several information-theoretic limits are also provided, including the Gaussian input Shannon limits, the CM limits and the BICM limits constrained by different constellations with/without rotation. These limits are derived via the AMI analysis shown in Section II.

The BER performance is shown in Fig. 9. As shown in this figure, the proposed BICM-ID-SSD systems are only about 0.7 to 0.8 dB from the CM limits at the BER of 10^{-5}, either for 16 or 64QAM, either with or without rotation. BICM systems with the DVB-T2 LDPC code also exhibit excellent performance that are only about 1.0 to 1.2 dB away from the BICM limits associated with 16/64QAM under Rayleigh fading channels. Since the CM limits are much lower than the BICM limits, the proposed BICM-ID-SSD systems can easily outperform these BICM systems. For example, the proposed BICM-ID-SSD systems with 16/64QAM and rotation angle of 45° are 0.8 to 1.0 dB better than the reference BICM systems, at the BER of 10^{-5}. Furthermore, the rotation gains are about 0.3/0.2 dB for 16/64QAM at the code rate of 1/2, which matches well with the AMI analysis shown in Fig. 3. Traditional BICM-ID systems without SSD display worse performance than the proposed BICM-ID-SSD systems. This is because the demapper's EXIT curves in traditional BICM-ID systems can not match well with the outer decoder's inverted EXIT curve simultaneously under different channels, as shown in Fig. 7. Therefore, it verifies that SSD plays a very important role for a BICM-ID-SSD system to hold simultaneously excellent performance under both AWGN and fading channels.

V. Conclusions

This paper investigated the issues of combining SSD with CMSs. A new criterion for determining the optimal rotation angle in SSD via maximizing the AMI is proposed. For optimal CMSs with SSD such as BICM-ID-SSD, or suboptimal independent CMSs such as BICM-SSD with Gray labeled square QAM, several important properties are obtained, based on which the optimization work is significantly simplified. For BICM-ID-SSD, the rotation angle of 45° is found optimal or near-optimal for square QAM at low to moderate code rates over Rayleigh fading channels. For BICM-SSD, it is interesting that the constellation rotation only has positive effect for low-order constellations (4QAM) at moderate to high code rates, but negative for high-order constellations (16/64QAM) at code rates of practical interests, or for low-order constellations at low code rates, over Rayleigh fading channels.

The tool of EXIT chart is used to help the BICM-ID-SSD system design. The problem that the demapper's EXIT curve exhibits different slopes under different channels can be mitigated by combining with SSD, which makes the proposed BICM-ID-SSD systems hold simultaneously excellent performance under both AWGN and fading channels. Simulation results show that the performance of the proposed BICM-ID-SSD systems are only about 0.7 to 0.8 dB from the CM limits, and more than 1.0 dB better than the reference BICM systems with DVB-T2 LDPC codes, at the code rate of $1/2$ and the BER of 10^{-5}, under Rayleigh fading channels.

Appendix: Proof of Lemma 2

We begin with the first item in Lemma 2 that the demapper's EXIT curve under fading channels has a smaller slope than that under AWGN channel. This statement is equivalent to that, for a given constellation and two different SNRs γ_0 and γ_1, if $I_{\text{CM}}(\gamma_0) = \mathbb{E}_\Lambda[I_{\text{CM}}(\Lambda^2\gamma_1)]$, the start point of the demapper's EXIT curve under AWGN channel is bellow that under fading channels, i.e., $I_{\text{BICM}}(\gamma_0) \leq \mathbb{E}_\Lambda[I_{\text{BICM}}(\Lambda^2\gamma_1)]$, where I_{CM} and I_{BICM} denote the CM-AMI and BICM-AMI functions over AWGN channels associated with the same constellation labeling. An intuitive proof is given as follows.

First, $I_{\text{CM}}(\gamma)$ and $I_{\text{BICM}}(\gamma)$ should appear like that shown in Fig. 10, because both of them are concave functions (all AMI cost functions should be concave [23]), $I_{\text{BICM}}(0) = I_{\text{CM}}(0) = 0$, $I_{\text{BICM}}(\gamma) \leq I_{\text{CM}}(\gamma)$, and $I_{\text{BICM}}(\gamma) \to I_{\text{CM}}(\gamma) \to m$

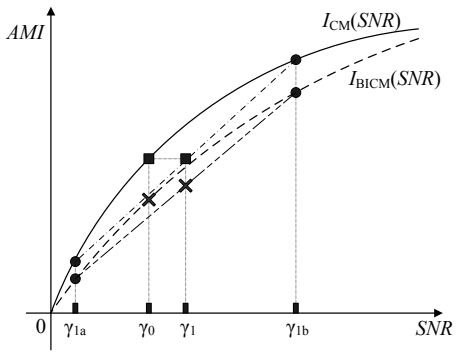

Fig. 10. The *relative concavity* of the CM-AMI function with respect to the BICM-AMI function.

as $\gamma \to \infty$. This implies that $I_{\mathrm{CM}}(\gamma)$ should be "more concave" than $I_{\mathrm{BICM}}(\gamma)$. Second, thinking about a fading channel with two states. When the overall SNR is γ_1, each state gets the SNR of γ_{1a} and γ_{1b} with equiprobability, i.e., $\gamma_1 = \frac{1}{2}(\gamma_{1a} + \gamma_{1b})$. This fading channel achieves the CM-AMI of $\frac{1}{2}(I_{\mathrm{CM}}(\gamma_{1a}) + I_{\mathrm{CM}}(\gamma_{1b}))$. To achieve the same CM-AMI, the AWGN channel only needs the SNR of γ_0 that $I_{\mathrm{CM}}(\gamma_0) = \frac{1}{2}(I_{\mathrm{CM}}(\gamma_{1a}) + I_{\mathrm{CM}}(\gamma_{1b}))$. Now we need to show that $I_{\mathrm{BICM}}(\gamma_0) \le \frac{1}{2}(I_{\mathrm{BICM}}(\gamma_{1a}) + I_{\mathrm{BICM}}(\gamma_{1b}))$. In fact, the "more concavity" can be mathematically described as *relative concavity*, which could be defined as follows [40]. Let g: $\mathbb{R} \to \mathbb{R}$ be strictly increasing on (a, b). f is said *relative concave* with respect to g if and only if for any three points $x_0 < y < x_1$ in the interval (a, b), we have

$$\frac{f(y) - f(x_0)}{f(x_1) - f(x_0)} \ge \frac{g(y) - g(x_0)}{g(x_1) - g(x_0)}.$$

Simply by assigning $g(x) = x$, the relative concave functions are degenerated to traditional concave functions. Now, by letting $g = I_{\mathrm{BICM}}, f = I_{\mathrm{CM}}, a = 0, b = +\infty$ and $x_0 = \gamma_{1a}, x_1 = \gamma_{1b}$, $y = \gamma_0$ (easy to verify that $\gamma_{1a} < \gamma_0 < \gamma_{1b}$), we have $I_{\mathrm{BICM}}(\gamma_0) \le \frac{1}{2}(I_{\mathrm{BICM}}(\gamma_{1a}) + I_{\mathrm{BICM}}(\gamma_{1b}))$. Note that this proof is based on the fact that I_{CM} is *relative concave* with respect to I_{BICM} according to the intuitive observations without a strict proof.

We now show the second item in Lemma 2. We first show that the extrinsic information of the demapper's output with perfect *a priori* information input in the case using SSD IE_1' is higher than that in the traditional case without SSD IE_1, under the same channel with the same SNR and constellation labeling. If perfect *a priori* information is available, to demap a certain bit, the other bits are exactly known. Therefore, it turns to be BPSK demapping. Let $A(x_{I0}, x_{Q0})$ represent the possible transmitted signal point corresponding to the bit being 0, and $B(x_{I1}, x_{Q1})$ being 1. These two points are moved to $A'(\Lambda_I x_{I0}, \Lambda_Q x_{Q0})$ and $B'(\Lambda_I x_{I1}, \Lambda_Q x_{Q1})$ at the receiver side, respectively, due to fading and the I/Q interleaving. By denoting $d = |AB|$ as the distance between A and B, it is easy to find that $d' = |A'B'| = \Gamma d$ where $\Gamma = \sqrt{\Lambda_I^2 \cos^2\varphi + \Lambda_Q^2 \sin^2\varphi}$ represents the equivalent fading coefficient, and φ denotes the angle between AB and the I-axis. The extrinsic information of the demapper's output of this bit in the traditional case without I/Q interleaving, denoted by IE_1^k, can be written as $IE_1^k = I(X; Y|\Lambda)$. With I/Q interleaving, such extrinsic information is changed to $IE_1'^k = I(X; Y|\Gamma)$. The goal now is to show that $I(X; Y|\Gamma) \ge I(X; Y|\Lambda)$. In fact, $I(X; Y|\Gamma) = \mathbb{E}\left[I_B(\Gamma^2\gamma)\right]$ and $I(X; Y|\Lambda) = \mathbb{E}\left[I_B(\Lambda^2\gamma)\right]$, where $I_B(\gamma)$ represents the CM-AMI (also the BICM-AMI) of the binary input AWGN channel with the SNR $\gamma = E_s^b/N_0$ whereby E_s^b denotes the equivalent BPSK symbol energy. Since $I_B(\gamma)$ is a concave function, the following inequality can be shown that

$$\begin{aligned} I(X; Y|\Gamma) &= \mathbb{E}\left[I_B(\Gamma^2\gamma)\right] \\ &= \mathbb{E}\left[I_B(\cos^2\varphi\Lambda_I^2\gamma + \sin^2\varphi\Lambda_Q^2\gamma)\right] \\ &\ge \cos^2\varphi\mathbb{E}\left[I_B(\Lambda_I^2\gamma)\right] + \sin^2\varphi\mathbb{E}\left[I_B(\Lambda_Q^2\gamma)\right] \\ &= \mathbb{E}\left[I_B(\Lambda^2\gamma)\right] = I(X; Y|\Lambda), \end{aligned}$$

where Λ represents the fading coefficient having the same distribution as Λ_I or Λ_Q. For any given bit, it has been shown that $IE_1'^k \ge IE_1^k$. It is easy to find that over all m bit,

$$IE_1' = \frac{1}{m}\sum_{k=0}^{m-1} IE_1'^k \ge \frac{1}{m}\sum_{k=0}^{m-1} IE_1^k = IE_1. \tag{9}$$

Therefore, higher extrinsic information of the demapper's output with perfect *a priori* information input can be gained via I/Q interleaving. It is worth mentioning that if $\Lambda_I \equiv \Lambda_Q$, i.e., without I/Q interleaving or over AWGN channels, the distance becomes $d' = \Lambda d$. If the binary signal set lies vertically ($\varphi = \pi/2$) or horizontally ($\varphi = 0$), we also have $d' = \Lambda d$. In these two cases, IE_1 with or without I/Q interleaving shall be exactly the same.

Since we already obtain a higher IE_1 by using SSD, if the area under the demapper's EXIT curve keeps the same or becomes smaller, a larger slope is achieved. If such area becomes larger, and the slope maintains or becomes smaller, then the curve with SSD will be above the original one at any IA point. Under either situation of the above two, we benefit from using SSD.

Based on the above proof, it has been shown that the equivalent channel turns to be a better fading channel by coordinate interleaving, when perfect *a priori* information is available at the demapper. If the constellation has N dimensions, then the equivalent channel fading coefficient $\Gamma^2 = \sum_{l=1}^{N} w_l\Lambda_l^2$ where w_ls denote the weights with $\sum_{l=1}^{N} w_l = 1$. Clearly it is easy to find that $\Gamma^2 \to 1$ as $N \to \infty$ according to the large number law, which implies that the equivalent channel is turned to be an AWGN channel. This result matches well with those derived in [13], [14], but yet more general that is suitable for any fading channel. Therefore, the terminal point of the demapper's EXIT curve with extremely high dimensional constellations and SSD under fading channels can approach that under AWGN channel, but the CM-AMI under the fading channels is obviously smaller than that under AWGN channel at the same SNR. Thereby, the demapper's EXIT curve under fading channels with SSD would hold a larger slope. However, approximately the same slope is highly required. Therefore, we should not use a very high-dimensional constellation. Another question that arises is how many dimensions are enough. No general answers have been obtained yet, but as shown in Fig. 4 and Fig. 7, 2-dimensional constellations are already acceptable under Rayleigh channels.

REFERENCES

[1] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1453–1467, July 1998.

[2] G. Taricco and E. Viterbo, "Performance of component interleaved signal sets for fading channels," *Electron. Lett.*, vol. 32, no. 13, pp. 1170–1172, Apr. 1996.

[3] M. N. Khormuji, U. H. Rizvi, G. J. Janssen, and S. B. Slimane, "Rotation optimization for MPSK/MQAM signal constellations over Rayleigh fading channels," in *Proc. IEEE ICCS*, Oct. 2006, pp. 1–5.

[4] N. F. Kiyani and J. H. Weber, "Performance analysis of a partially coherent system using constellation rotation and coordinate interleaving," in *Proc. IEEE Globecom*, Dec. 2008, pp. 1–5.

[5] ——, "Performance analysis of a system using coordinate interleaving and constellation rotation in Rayleigh fading channels," in *Proc. IEEE VTC Fall*, Sep. 2008, pp. 1–5.

[6] A. Chindapol and J. A. Ritcey, "Bit-interleaved coded modulation with signal space diversity in Rayleigh fading," in *Proc. 33rd Asilomar Conf. Signals, Systems, Computers*, 1999, pp. 1003–1007.

[7] N. F. Kiyani, U. H. Rizvi, J. H. Weber, and G. J. Janssen, "Optimized rotations for LDPC-coded MPSK constellations with signal space diversity," in *Proc. IEEE WCNC*, Mar. 2007, pp. 677–681.

[8] C. A. Nour and C. Douillard, "Improving BICM performance of QAM constellations for broadcasting applications," in *Proc. 5th Int. Symp. Turbo Codes and Related Topics*, 2008, pp. 55–60.

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

[10] Y. Li, X.-G. Xia, and G. Wang, "Simple iterative methods to exploit the signal-space diversity," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 32–38, Jan. 2005.

[11] N. F. Kiyani and J. H. Weber, "OFDM with BICM-ID and rotated MPSK constellations and signal space diversity," in *Proc. IEEE Symp. on Commun. and Veh. Technology*, Nov. 2007, pp. 1–4.

[12] ——, "EXIT chart analysis of iterative demodulation and decoding of MPSK constellations with signal space diversity," *J. Commun.*, vol. 3, no. 3, pp. 43–50, July 2008.

[13] N. H. Tran, H. H. Nguyen, and T. Le-Ngoc, "Performance of BICM-ID with signal space diversity," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1732–1742, May 2007.

[14] ——, "BICM-ID with signal space diversity over cascaded Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 56, no. 10, pp. 1561–1568, Oct. 2008.

[15] ——, "Performance analysis and design criteria of BICM-ID with signal space diversity for Keyhole Nakagami-m fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1592–1602, Apr. 2009.

[16] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.

[17] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding using soft feedback," *Electron. Lett.*, vol. 34, no. 10, pp. 942–943, May 1998.

[18] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Globecom*, 1998, pp. 579–584.

[19] Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2), ETSI Std. DVB Document A122, June 2008.

[20] L. Vangelista, N. Benvenuto, S. Tomasin, C. Nokes, J. Stott, A. Filippi, M. Vlot, V. Mignone, and A. Morello, "Key technologies for next-generation terrestrial digital television standard DVB-T2," *IEEE Commun. Mag.*, vol. 47, pp. 146–153, Oct. 2009.

[21] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[22] S. Pfletschinger and F. Sanzi, "Error floor removal for bit-interleaved coded modulation with iterative detection," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3174–3181, Nov. 2006.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[24] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[25] S. Y. Goff, "Signal constellations for bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 307–313, Jan. 2003.

[26] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

[27] F. Schreckenbach, "Iterative decoding of bit-interleaved coded modulation," Ph.D. dissertation, University of Stuttgart, Sep. 2007.

[28] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE ICC*, May 1993, pp. 1064–1070.

[29] F. Schreckenbach and G. Bauch, "EXIT charts for iteratively decoded multilevel modulation," in *Proc. EUSIPCO*, 2004.

[30] Q. Xie, K. Peng, F. Yang, and Z. Wang, "BICM-ID systems with signal space diversity over Rayleigh fading channels," in *Proc. IEEE Globecom*, Dec. 2010.

[31] 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, 2002.

[32] J. Tan and G. L. Stuber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 662–672, Mar. 2005.

[33] N. H. Tran and H. H. Nguyen, "A novel multi-demensional mapping of 8-PSK for BICM-ID," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1133–1142, Mar. 2007.

[34] R. Y. Tee, R. G. Maunder, and L. Hanzo, "EXIT-chart aided near-capacity irregular bit-interleaved coded modulation design," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 32–37, Jan. 2009.

[35] A. Sezgin and E. A. Jorswieck, "Impact of the mapping strategy on the performance of APP decoded spacectime block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4685–4690, Dec. 2005.

[36] F. Schreckenbach, N. Gortz, J. Hagenauer, and G. Bauch, "Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding," in *Proc. IEEE Globecom*, Dec. 2003, pp. 3316–3320.

[37] ——, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[38] D. J. C. MacKay, "Good error correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, pp. 399–431, Mar. 1999.

[39] S. Dolinar and D. Divsalar, "Weight distribution for turbo codes using random and nonrandom permutations," TDA Progress Rep. 42-122, JPL, pp. 56–65, Aug. 1995.

[40] J. A. Palmer, "Relative convexity," Technical Report, ECE Dept., UCSD, 2003.

Qiuliang Xie was born in Hunan, China. He received his B.Eng degree in School of Telecommunication Engineering from Beijing University of Post and Telecommunications, Beijing, China, in 2006. He is pursuing his Ph.D degree in the department of Electronic Engineering since 2006, Tsinghua University. His main research interests focus on the physical layer broadband wireless communication and broadcasting technologies, especially the information theory, channel coding & coded modulation, MIMO, relay, multi-user detection, space-time coding as well as channel estimation techniques.

Jian Song received his B. Eng and Ph.D degrees from the Department of Electrical Engineering, Tsinghua University, China in 1990 and 1995, respectively. He worked for Tsinghua University upon graduation and was Postdoctoral research fellow at Chinese University of Hong Kong and University of Waterloo, Canada in 1996 and 1997, respectively. After that, Dr. Song was with Hughes Network Systems in USA before joining the faculty team at Tsinghua in 2005 as a full professor. He is the Director of DTV Technology R&D Center and the deputy chairman of Department of Electronic Engineering. Dr. Song is the associate editor of IEEE Transactions on Broadcasting and the Chairman as well as founding person of IEEE BTS Beijing Chapter. Dr. Song's primary research interest is in the physical layer transmission technologies and he has been working in different areas, such as fiber-optic, satellite and wireless communications, as well as the powerline communications. His current research interest is in the digital TV broadcasting area. Dr. Song has published more than 90 peer-reviewed journal and conference papers, holds two US and ten Chinese patents. He is also the co-author of one book in DTV area and translated one book in powerline communications area.

Kewu Peng was born in Hefei, China. He received the B.E. degree in Electronics Engineering from Hefei University of Technology in 1993, the M.E. degree in Electronics Engineering from Tsinghua University in 1996, and the Ph.D. degree in Electrical and Computer Engineering from the University of Minnesota, Minneapolis, in 2003. He was a Researcher and Lecturer with the Department of Electronic Engineering, Tsinghua University, from Aug 1996 to Aug 1999. Since Jan 2005, he joins the Digital Television Research Center at Tsinghua University as a Research Staff, Assistant Researcher, and Associate Professor('09). His research interests include mobile/wireless communications, digital terrestrial/television broadcasting, and image/video transmission.

Fang Yang received his B.S.E. and Ph.D. degree in 2005 and 2009, respectively, from the Department of Electronic Engineering in Tsinghua University, Beijing China. Currently, he is working as an assistant researcher at the DTV Technology R&D Center, Tsinghua University. His research interests lie in the field of channel estimation and interference cancellation for digital wireless communication system, space-time coding and diversity techniques, as well as the training sequence design.

Zhaocheng Wang received his B.S.E., M.S.E. and Ph.D. degrees from Tsinghua University in 1991, 1993 and 1996 respectively. He was a Post Doctoral Fellow with Nanyang Technological University (NTU) in Singapore from 1996 to 1997. After that, he was with OKI Techno Centre (Singapore) Pte. Ltd. from 1997 to 1999, firstly as a research engineer and then as a senior engineer. From 1999 to 2009, he worked at SONY Deutschland GmbH, firstly as a senior engineer and then as a principal engineer. Since April 2009, he has been with Department of Electronic Engineering at Tsinghua University as a Full Professor. His general research interests include wireless communications, digital video broadcasting and signal processing. Within these areas, he held 18 US/EU patents and published more than 40 technical papers.