Design and Analysis of BICM-ID for Two-Way Relay Channels With Physical-Layer Network Coding

Zichao Sun, Li Chen, *Senior Member, IEEE*, Xiaojun Yuan, *Member, IEEE*, and Yushan Yakufu

***Abstract*—This paper proposes a general iterative demapping–decoding approach for two-way relay (TWR) communications with physical-layer network coding, where the relay node performs decode-and-forward. The bit-interleaved coded modulation (BICM) with high-order quadrature amplitude modulation is considered. The iterative decoding, namely the BICM-ID for TWR channels, is introduced. Based on a joint trellis, BICM-ID estimates the end nodes' coded bits, leading to an estimation of the transmitted symbol pair probability, which is crucial for achieving iterative decoding gains in bit error rate (BER). The end nodes' constellation design criterion is proposed to optimize the BICM-ID performance over the TWR fading channel. The EXtrinsic Information Transfer analysis is also performed, showing the iterative system's convergence behavior and providing information on code design. Our simulation results show the proposed BICM-ID scheme can achieve significant performance gains, yielding an asymptotic tendency of approaching the cut-set BER performance limit.**

***Index Terms*—BICM, convolutional codes, constellation design, EXIT analysis, spectrum efficiency, TWR communications.**

I. Introduction

NETWORK coding (NC) [1] is a celebrated technique for improving information throughput in a communication network. Two-way relay (TWR) communications in which two end nodes exchange information through a relay node is where NC plays an important role. Further assisted by physical layer NC (PNC) [2]–[5], information throughput of TWR communications can be further improved [4]. With PNC, relay node can decode, e.g., the exclusive-or (XOR) information of the end nodes, based on the superimposed received symbols. This XOR information will be further encoded and broadcast to the end nodes.

Manuscript received December 8, 2016; revised September 19, 2017; accepted September 20, 2017. Date of publication September 27, 2017; date of current version November 10, 2017. This work was supported in part by the National Natural Science Foundation of China under Projects 61372079, 61671486, and 61471241, and in part by the Chinese Recruitment Program for Global Young Experts. The review of this paper was coordinated by Dr. R. D. Souza. (*Corresponding author: Li Chen.)*

Z. Sun and Y. Yakufu are with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China (e-mail: sunzch@mail2.sysu.edu.cn; yakufu@mail2.sysu.edu.cn).

L. Chen is with the School of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou 510006 China (e-mail: chenli55@mail.sysu.edu.cn).

X. Yuan is with the National Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 610051, China (e-mail: xjyuan@uestc.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2017.2757151

The capacity region of the end nodes in the PNC assisted TWR Gaussian channel has been characterized by [6], [7]. In approaching the capacity, lattice code has been considered in [6]–[10] with an attempt to harness the multiple access (MAC) phase interference. The design of repeat-accumulate (RA) codes, irregular RA (IRA) codes and low-density parity-check (LDPC) codes for TWR Gaussian channel have been investigated in [5], [11] and [12], respectively. The above mentioned work employed BPSK modulation limiting the spectrum efficiency. It has been aware that the demapping at the relay node becomes problematic if high order modulation schemes are employed. For example, over the TWR Gaussian channel, one superimposed received symbol may correspond to multiple XOR coded bit permutations. Addressing this challenge, an iterative demapping-decoding approach for the bit-interleaved coded modulation (BICM) coded TWR communication systems has been proposed [13]. However, it has been realized that decoding performance can be further improved by a more comprehensive demapping at the relay. This was indicated by the earlier work in multi-user detection (MUD) [14]–[16] where demapping is performed based on the superimposed signal of multiple users. It requires coherent coding and modulation schemes to be employed by all users.

The phase and symbol arrival-time asynchrony in PNC assisted TWR communications has recently been addressed by [17] in which a three-layer decoding framework at the relay was proposed. On the other aspect, constellation design for an uncoded TWR communication system has been considered in [18], [19]. In particular, the closest-neighbor clustering constellation design was proposed by [18] to better denoise the MAC phase interference, while the high order quadrature amplitude modulation (QAM) constellation was designed utilizing Latin squares by [19]. Recently, the channel coded TWR communications using high order modulations have been considered in [20]–[23]. The iterative noncoherent receiver design for TWR communications was considered by [20]. The side information of high order modulations was exploited to distinguish the two end nodes' signals by [21]. The design of QPSK for BICM-ID over the TWR channels was considered by [22]. Our recent work has noticed that the demapping challenge in using high order modulation schemes can also be better addressed by iterative demapping-decoding.

This paper proposes the BICM coded TWR communication that employs the energy efficient high order QAMs. To improve the decoding performance at relay, the iterative demapping-decoding approach based on a joint Viterbi trellis, namely the

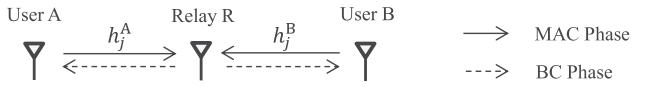

Fig. 1. PNC assisted TWR communications.

BICM-ID, is proposed. The relay can completely decode the end nodes' information bits. Compared with our earlier work [13], using a joint trellis not only improves the decoding performance, but also gives a greater design freedom for the end nodes since they can employ different convolutional codes. For example, one can deploy a stronger code if its uplink channel to the relay exhibits a greater fading attenuation. This distinguishes our contribution to the work in MUD where coherent coding scheme is required for all users. In our proposal, the demapper obtains channel observations with knowledge of the end nodes' mapping constellations and perfect channel state information (CSI) [1] of the uplink channels. By iteratively updating the transmitted symbol pair probabilities, the demapper can generate more accurate *a priori* information of the end nodes' coded bits. Consequently, the decoder can generate more accurate estimations for the end nodes' information and coded bits. In order to optimize the decoding performance, design criterion of the end nodes' mapping constellation is proposed under the content of TWR fading channel. Furthermore, the EXtrinsic Information Transfer (EXIT) analysis [24] is carried out to reveal the iterative decoding convergence over the TWR fading channel. Using the EXIT chart, we can also better design channel codes to maximize the decoding performance at relay. Finally, our simulation results show that significant decoding gains can be achieved by iterations. They approach the cut-set performance limit in the high signal-to-noise ratio (SNR) region.

The rest of this paper is organized as follows. Section II presents the signal model of the PNC assisted TWR communications with BICM and BICM-ID. Section III proposes the BICM-ID scheme. Section IV introduces the constellation design criterion for the TWR fading channel. The EXIT analysis is performed in Section V. Section VI shows our simulation results, and finally Section VII concludes the paper.

II. Signal Model

Fig. 1 shows the PNC assisted TWR communication model which consists of two end nodes, A and B, and a relay node, R. There is no direct link between A and B. They have to exchange message through R. In this paper, we assume that the channel is memoryless, CSI is perfectly known by the relay and perfect frame synchronization.

In the MAC phase, A and B simultaneously transmit their codewords to R. Receiving the noisy superimposed signal, R decodes the transmitted messages of the end nodes. It further generates the XOR NC message and encodes it for broadcasting (BC) to the end nodes. Each end node can then recover the other's message if it can decode the NC message.

Fig. 2 shows the proposed BICM coded TWR communication system and the BICM-ID that is performed by the relay, where the key operations are highlighted by equation numbers. For simplicity, we assume that both end nodes employ convolutional codes of the same rate and modulation schemes of the same order. In particular, M-ary QAMs are employed by the end nodes with order $m = \log_2 M$. Let $\mathbf{u}_\xi = [u_1^\xi, u_2^\xi, \ldots, u_t^\xi, \ldots, u_l^\xi]$ denote the binary message of end node ξ, where l is the message length and $\xi \in \{\mathrm{A}, \mathrm{B}\}$. Assume that the end nodes employ convolutional codes of rate r. The encoder generates the codeword $\mathbf{c}_\xi = [c_1^\xi, c_2^\xi, \ldots, c_{t'}^\xi, \ldots, c_{\frac{l}{r}}^\xi]$. In BICM, the serial-to-parallel (S/P) converter generates m parallel coded bit sequences $\{\mathbf{c}^{\xi,1}, \mathbf{c}^{\xi,2}, \ldots, \mathbf{c}^{\xi,m}\} = \{\{c_{j,1}^\xi\}, \{c_{j,2}^\xi\}, \ldots, \{c_{j,m}^\xi\}\}$, where $j = 1, 2, \ldots, \frac{l}{rm}$ [2]. They are then randomly interleaved, yielding m interleaved coded bit sequences $\{\mathbf{v}^{\xi,1}, \mathbf{v}^{\xi,2}, \ldots, \mathbf{v}^{\xi,m}\} = \{\{v_{j,1}^\xi\}, \{v_{j,2}^\xi\}, \ldots, \{v_{j,m}^\xi\}\}$ [3]. Note that the two end nodes can employ different interleaving patterns $(\Pi_1^\xi, \Pi_2^\xi, \ldots, \Pi_m^\xi)$ for the m sequences. Afterwards, every m interleaved coded bits $\mathbf{v}_j^\xi = [v_{j,1}^\xi, v_{j,2}^\xi, \ldots, v_{j,m}^\xi]$ are mapped to a QAM symbol for transmission as

$$x_j^\xi = \Lambda_\xi\left(\mathbf{v}_j^\xi\right), \tag{1}$$

where $\Lambda_\xi(\cdot)$ denotes the mapping function of node ξ and $\mathbb{E}[||x_j^\xi||^2] = 1$ is maintained. Inversely, $\Lambda_\xi^{-1}(\cdot)$ denotes the demapping function. Let χ_ξ denote the signal constellation employed by node ξ and $x_j^\xi \in \chi_\xi$. In this work, we assume $\log_2|\chi_\mathrm{A}| = \log_2|\chi_\mathrm{B}|$ so that constellations of the end nodes have the same cardinality.

In the MAC phase, the signal received by R is

$$y_j = h_j^\mathrm{A} x_j^\mathrm{A} + h_j^\mathrm{B} x_j^\mathrm{B} + z_j, \tag{2}$$

where h_j^ξ denotes the complex fading coefficients of the ξ-to-R channel. They are independent and Rayleigh distributed with fading gains normalized as $\mathbb{E}[||h_j^\mathrm{A}||^2] = \mathbb{E}[||h_j^\mathrm{B}||^2] = 1$. Note that when $h_j^\mathrm{A} = h_j^\mathrm{B} = 1$, $\forall j$, the above equation depicts R's received signal over the TWR Gaussian channel. Although Rayleigh fading channel is adopted, our BICM-ID scheme can also be applied in other channel model, e.g., the Rician fading channel model. z_j is the complex additive white Gaussian noise (AWGN) that is observed by R with a variance of $N_0/2$ per dimension. The SNR is defined as $\rho \doteq E_b/N_0$, where E_b is the average energy per end node's message bit.

After the MAC transmission, R receives $\mathbf{y} = [y_1, y_2, \ldots, y_{\frac{l}{rm}}] \in \mathbb{C}^{\frac{l}{rm}}$ and employs BICM-ID to decode the end nodes' messages. Let us denote codeword sequence of the two end nodes as

$$\mathbf{c}_\mathrm{AB} = \Big[c_1^\mathrm{A}, c_2^\mathrm{A}, \ldots, c_n^\mathrm{A}, c_1^\mathrm{B}, c_2^\mathrm{B}, \ldots, c_n^\mathrm{B}, \ldots\ldots, c_{\frac{l}{r}-n+1}^\mathrm{A}, c_{\frac{l}{r}-n+2}^\mathrm{A}, \ldots, c_{\frac{l}{r}}^\mathrm{A}, c_{\frac{l}{r}-n+1}^\mathrm{B}, c_{\frac{l}{r}-n+2}^\mathrm{B}, \ldots, c_{\frac{l}{r}}^\mathrm{B}\Big], \tag{3}$$

where n denotes the number of coded bits produced by each trellis transition. Relay attempts to decode the messages of

[1] In practice, CSI can be obtained by pilot symbol training.

[2] It is assumed that l is divisible by rm.

[3] Note that the m interleavers can also be replaced by a single interleaver for the whole coded bit sequence.

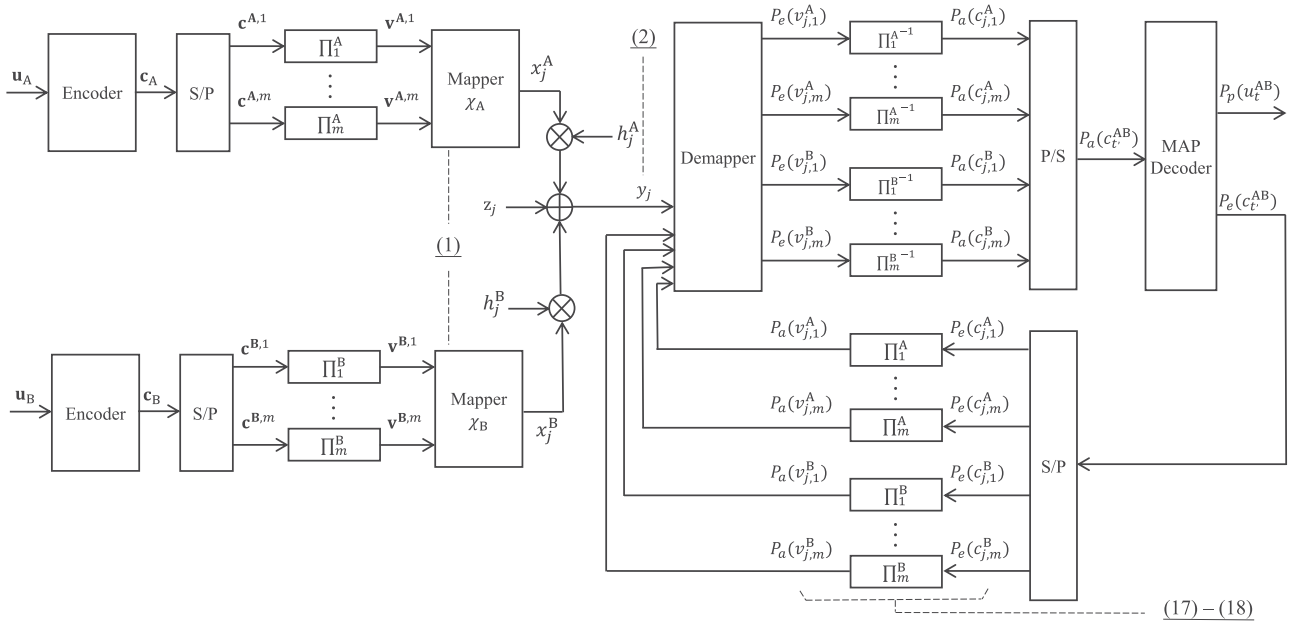

Fig. 2. The BICM coded TWR communication system and BICM-ID at R.

A and B

$$\mathbf{u}_{\mathrm{AB}} = \left[u_1^{\mathrm{A}}, u_2^{\mathrm{A}}, \ldots, u_k^{\mathrm{A}}, u_1^{\mathrm{B}}, u_2^{\mathrm{B}}, \ldots, u_k^{\mathrm{B}}, \ldots\ldots, u_{l-k+1}^{\mathrm{A}}, u_{l-k+2}^{\mathrm{A}}, \ldots, u_l^{\mathrm{A}}, u_{l-k+1}^{\mathrm{B}}, u_{l-k+2}^{\mathrm{B}}, \ldots, u_l^{\mathrm{B}}\right], \quad (4)$$

where k denotes the number of message bits that trigger each trellis transition. If $\mathbf{u}_{\mathrm{AB}}$ can be decoded, R generates the NC message by

$$\mathbf{u}_N = \left[u_1^{\mathrm{A}} \oplus u_1^{\mathrm{B}}, u_2^{\mathrm{A}} \oplus u_2^{\mathrm{B}}, \ldots, u_k^{\mathrm{A}} \oplus u_k^{\mathrm{B}}, \ldots, u_l^{\mathrm{A}} \oplus u_l^{\mathrm{B}}\right]. \quad (5)$$

Finally, the NC message will be encoded and broadcast to A and B during the BC phase. In this paper, we focus on decoding $\mathbf{u}_{\mathrm{AB}}$ based on $\mathbf{y}$.

III. BICM-ID for TWR Channels

This section introduces the joint trellis based BICM-ID scheme for the TWR channels. Block diagram of the proposal has been shown by Fig. 2.

A. Notations

To underpin the iterative demapping-decoding, the following notations need to be introduced.

1) Let P_a, P_p and P_e denote the *a priori*, the *a posteriori* and the extrinsic probabilities, respectively;
2) Let $\mathbf{v}_j^{\mathrm{AB}} = [\mathbf{v}_j^{\mathrm{A}}, \mathbf{v}_j^{\mathrm{B}}] = [v_{j,1}^{\mathrm{A}}, v_{j,2}^{\mathrm{A}}, \ldots, v_{j,m}^{\mathrm{A}}, v_{j,1}^{\mathrm{B}}, v_{j,2}^{\mathrm{B}}, \ldots, v_{j,m}^{\mathrm{B}}]$ denote the set of the end nodes' interleaved coded bits, where $\mathbf{v}_j^{\mathrm{A}}$ and $\mathbf{v}_j^{\mathrm{B}}$ are mapped to the transmitted symbols of A and B, respectively. Furthermore, let $\mathbf{v}_{j,i}^{\mathrm{AB}}$ denote the ith bit of $\mathbf{v}_j^{\mathrm{AB}}$ where $i = 1, 2, \ldots, 2m$. In particular,

$$\mathbf{v}_{j,i}^{\mathrm{AB}} = \begin{cases} v_{j,i}^{\mathrm{A}}, & \text{if } 1 \le i \le m, \\ v_{j,i-m}^{\mathrm{B}}, & \text{if } m+1 \le i \le 2m; \end{cases} \quad (6)$$

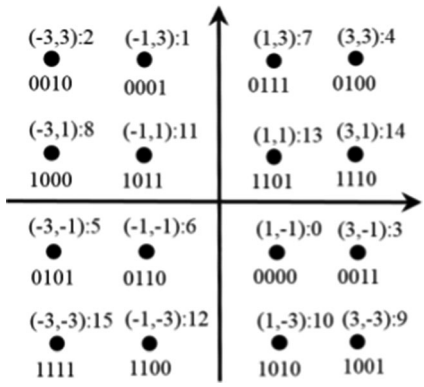

Fig. 3. MSEW 16QAM constellation.

3) Let s_ϱ^{A} and s_σ^{B} denote the constellation points that are chosen from χ_{A} and χ_{B}, respectively, and $\varrho, \sigma \in \{0, 1, \ldots, M-1\}$. For example, Fig. 3 shows the maximum squared Euclidean weight (MSEW) 16 QAM constellation [26]. Its design is guided by the maximum likelihood (ML) decoding bound of BICM systems to achieve an improved asymptotic bit error rate (BER) performance. If it is employed by node A, we denote $s_0^{\mathrm{A}} = (1, -1)$, $s_1^{\mathrm{A}} = (-1, 3)$, and etc. Furthermore, let $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$ denote a transmitted symbol pair of the two end nodes. If both end nodes employ the MSEW 16QAM, we can have $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}}) = (s_7^{\mathrm{A}}, s_{15}^{\mathrm{B}}) = ((1, 3), (-3, -3))$. In this case, we know $\mathbf{v}_j^{\mathrm{A}} = [0\ 1\ 1\ 1]$ and $\mathbf{v}_j^{\mathrm{B}} = [1\ 1\ 1\ 1]$;
4) Let $\chi_{i,b}^{\xi}$ denote node ξ's constellation points whose demapping produces the ith bit being b, where $b \in \{0, 1\}$ and $i = 1, 2, \ldots, m$. Based on the above introduction, it is known that $\Lambda_{\mathrm{A}}^{-1}(s_\varrho^{\mathrm{A}}) = [v_{j,1}^{\mathrm{A}}, v_{j,2}^{\mathrm{A}}, \ldots, v_{j,m}^{\mathrm{A}}]$ and $\Lambda_{\mathrm{B}}^{-1}(s_\sigma^{\mathrm{B}}) = [v_{j,1}^{\mathrm{B}}, v_{j,2}^{\mathrm{B}}, \ldots, v_{j,m}^{\mathrm{B}}]$. We can denote

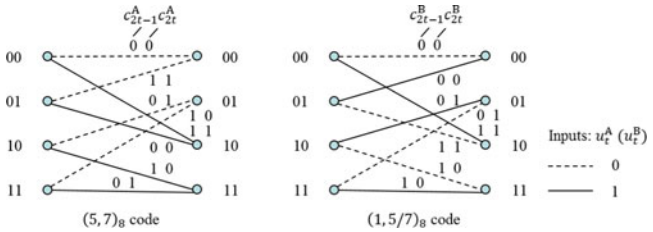

Fig. 4. Trellises of the $(5, 7)_8$ and the $(1, 5/7)_8$ convolutional codes.

$[\Lambda_{\mathrm{A}}^{-1}(s_{\varrho}^{\mathrm{A}})]_i = v_{j,i}^{\mathrm{A}}$ and $[\Lambda_{\mathrm{B}}^{-1}(s_{\sigma}^{\mathrm{B}})]_i = v_{j,i}^{\mathrm{B}}$. Consequently, $\chi_{i,b}^{\mathrm{A}}$ and $\chi_{i,b}^{\mathrm{B}}$ can be defined as

$$\chi_{i,b}^{\mathrm{A}} = \left\{ s_{\varrho}^{\mathrm{A}} | [\Lambda_{\mathrm{A}}^{-1}(s_{\varrho}^{\mathrm{A}})]_i = b \right\}, \tag{7}$$

$$\chi_{i,b}^{\mathrm{B}} = \left\{ s_{\sigma}^{\mathrm{B}} | [\Lambda_{\mathrm{B}}^{-1}(s_{\sigma}^{\mathrm{B}})]_i = b \right\}. \tag{8}$$

For example, if node A employs the MSEW 16QAM, we have $\chi_{1,1}^{\mathrm{A}} = \{s_8^{\mathrm{A}}, s_9^{\mathrm{A}}, s_{10}^{\mathrm{A}}, s_{11}^{\mathrm{A}}, s_{12}^{\mathrm{A}}, s_{13}^{\mathrm{A}}, s_{14}^{\mathrm{A}}, s_{15}^{\mathrm{A}}\}$;

5) Let $\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ denote the $2m$ interleaved coded bits that correspond to symbol pair $(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ as

$$\begin{aligned} \mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}) &= \left[\Lambda_{\mathrm{A}}^{-1}(s_{\varrho}^{\mathrm{A}}), \Lambda_{\mathrm{B}}^{-1}(s_{\sigma}^{\mathrm{B}}) \right] \\ &= \left[v_{j,1}^{\mathrm{A}}, v_{j,2}^{\mathrm{A}}, \ldots, v_{j,m}^{\mathrm{A}}, v_{j,1}^{\mathrm{B}}, v_{j,2}^{\mathrm{B}}, \ldots, v_{j,m}^{\mathrm{B}} \right]. \end{aligned} \tag{9}$$

Pursuant to the example in 3), we can denote $\mathbf{v}^{\mathrm{AB}}(s_7^{\mathrm{A}}, s_{15}^{\mathrm{B}}) = [0\,1\,1\,1\,1\,1\,1\,1]$. Furthermore, let $[\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})]_i$ denote its ith bit. Integrating (6), we have $[\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})]_i = v_{j,i}^{\mathrm{A}}$ if $1 \le i \le m$, and $[\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})]_i = v_{j,i-m}^{\mathrm{B}}$ if $m+1 \le i \le 2m$;

6) Finally, let $(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})_{i,b}$ denote the set of transmitted symbol pairs whose demapping produces the ith bit being b as

$$\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right)_{i,b} = \left\{ (s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}) \mid [\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})]_i = b \right\}. \tag{10}$$

Again, if both end nodes employ the MSEW 16QAM, we have $(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})_{1,1} = \{(s_8^{\mathrm{A}}, s_0^{\mathrm{B}}), (s_8^{\mathrm{A}}, s_1^{\mathrm{B}}), \ldots, (s_8^{\mathrm{A}}, s_{15}^{\mathrm{B}}), (s_9^{\mathrm{A}}, s_0^{\mathrm{B}}), (s_9^{\mathrm{A}}, s_1^{\mathrm{B}}), \ldots, (s_9^{\mathrm{A}}, s_{15}^{\mathrm{B}}), \ldots, (s_{15}^{\mathrm{A}}, s_0^{\mathrm{B}}), (s_{15}^{\mathrm{A}}, s_1^{\mathrm{B}}), \ldots, (s_{15}^{\mathrm{A}}, s_{15}^{\mathrm{B}})\}$, and $|(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})_{1,1}| = 128$.

B. Joint Trellis

With a joint trellis, we are able to estimate the two end nodes' messages $\mathbf{u}_{\mathrm{A}}$ and $\mathbf{u}_{\mathrm{B}}$ using the maximum *a posteriori* (MAP) decoding algorithm [27]. The joint trellis is cascaded by the two end nodes' trellises [28]. For example, Fig. 4 shows the trellises of the four states $(5, 7)_8$ and $(1, 5/7)_8$ convolutional codes where the encoder structures are notated in the octal form. They are rate half codes with $k = 1$ and $n = 2$. If the two codes are employed by nodes A and B, respectively, a message bit of u_t^{A} (u_t^{B}) generates two coded bits of $c_{2t-1}^{\mathrm{A}} c_{2t}^{\mathrm{A}}$ ($c_{2t-1}^{\mathrm{B}} c_{2t}^{\mathrm{B}}$) as indicated in the trellises. The joint trellis that is employed by the relay will have 16 states with $u_t^{\mathrm{A}} u_t^{\mathrm{B}}$ and $c_{2t-1}^{\mathrm{A}} c_{2t}^{\mathrm{A}} c_{2t-1}^{\mathrm{B}} c_{2t}^{\mathrm{B}}$ as the input and output of each trellis branch. In our proposal, the demapping will provide the *a priori* information for each of the end nodes' coded bits. They will form the branch metrics for the MAP decoding. The above description shows that the end nodes can also employ codes of different rate and different number of states.

C. BICM-ID

With the received symbol vector $\mathbf{y}$ and CSI, the demapper obtains the channel observations by

$$\Pr\left(y_j | s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) = \frac{1}{\pi N_0} \exp\left(-\frac{||y_j - h_j^{\mathrm{A}} s_{\varrho}^{\mathrm{A}} - h_j^{\mathrm{B}} s_{\sigma}^{\mathrm{B}}||^2}{N_0} \right). \tag{11}$$

With the M-ary QAM being employed by the end nodes, each received symbol y_j spins out M^2 different channel observations that correspond to M^2 possible transmitted symbol pairs $(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$. The demapper determines the *a posteriori* probabilities (APPs) of the end nodes' interleaved coded bit $\mathbf{v}_{j,i}^{\mathrm{AB}}$ by

$$\begin{aligned} P_p\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b | y_j\right) &= \mathcal{N}_1 \sum_{(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})_{i,b}} \Pr\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}} | y_j\right) \\ &= \mathcal{N}_1 \sum_{(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})_{i,b}} \Pr\left(y_j | s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) \cdot \Pr\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right), \end{aligned} \tag{12}$$

where $\mathcal{N}_1 = (\sum_{b \in \{0,1\}} P_p(\mathbf{v}_{j,i}^{\mathrm{AB}} = b | y_j))^{-1}$ is a normalization factor [4]. In the above equation, $\Pr(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ is the probability of symbols s_{ϱ}^{A} and s_{σ}^{B} being transmitted by the two end nodes, and it is called the *transmitted symbol pair probability*. Since messages $\mathbf{u}_{\mathrm{A}}$ and $\mathbf{u}_{\mathrm{B}}$ are independent, we have

$$\Pr\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) = \Pr\left(s_{\varrho}^{\mathrm{A}}\right) \cdot \Pr\left(s_{\sigma}^{\mathrm{B}}\right). \tag{13}$$

Using a joint trellis, we can determine $\Pr(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ by estimating $\Pr(s_{\varrho}^{\mathrm{A}})$ and $\Pr(s_{\sigma}^{\mathrm{B}})$. Since each permutation of $[v_{j,1}^{\mathrm{A}}, v_{j,2}^{\mathrm{A}}, \ldots, v_{j,m}^{\mathrm{A}}, v_{j,1}^{\mathrm{B}}, v_{j,2}^{\mathrm{B}}, \ldots, v_{j,m}^{\mathrm{B}}]$ corresponds to one transmitted symbol pair and interleaving enables the $2m$ bits of $\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ being independent, $\Pr(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ can be determined by

$$\begin{aligned} \Pr\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) &= \Pr\left(s_{\varrho}^{\mathrm{A}}\right) \cdot \Pr\left(s_{\sigma}^{\mathrm{B}}\right) \\ &= \prod_{i=1}^{m} P_a\left(\left[\Lambda_{\mathrm{A}}^{-1}\left(s_{\varrho}^{\mathrm{A}}\right) \right]_i \right) \cdot \prod_{i=1}^{m} P_a\left(\left[\Lambda_{\mathrm{B}}^{-1}\left(s_{\sigma}^{\mathrm{B}}\right) \right]_i \right). \end{aligned} \tag{14}$$

We can also denote $\Pr(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$ as

$$\Pr\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) = \prod_{i=1}^{2m} P_a\left(\left[\mathbf{v}^{\mathrm{AB}}\left(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}\right) \right]_i \right). \tag{15}$$

Now, let us denote $[\mathbf{v}^{\mathrm{AB}}(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})]_i = \mathbf{v}_{j,i}^{\mathrm{AB}}$. At the beginning of the iterative decoding, no knowledge of $\mathbf{v}_{j,i}^{\mathrm{AB}}$ is available and the demapper assumes $P_a(\mathbf{v}_{j,i}^{\mathrm{AB}} = 0) = P_a(\mathbf{v}_{j,i}^{\mathrm{AB}} = 1) = \frac{1}{2}$, and $\Pr(s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}}) = \frac{1}{2^{2m}} = \frac{1}{M^2}, \forall (s_{\varrho}^{\mathrm{A}}, s_{\sigma}^{\mathrm{B}})$. This is equivalent to assuming all the points of χ_{A} and χ_{B} have equal occurrence such that

[4] $\Pr(y_j)$ is omitted in (12) since it is neutralized by the normalization.

$\Pr(s_\varrho^{\mathrm{A}}) = \Pr(s_\sigma^{\mathrm{B}}) = \frac{1}{M}$. After the MAP decoding, more information on $\mathbf{v}_{j,i}^{\mathrm{AB}}$ will be available and they can be utilized to update $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$ as in (14). Furthermore, the extrinsic probabilities of $\mathbf{v}_{j,i}^{\mathrm{AB}}$ can be determined by

$$P_e\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b\right) = \mathcal{N}_2 \frac{P_p\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b | y_j\right)}{P_a\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b\right)}, \tag{16}$$

where $\mathcal{N}_2 = (\sum_{b\in\{0,1\}} P_e(\mathbf{v}_{j,i}^{\mathrm{AB}} = b))^{-1}$ is a normalization factor. Since $P_a(\mathbf{v}_{j,i}^{\mathrm{AB}} = b)$ can also be denoted as $P_a([\mathbf{v}^{\mathrm{AB}}(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})]_i = b)$, further based on (12)–(15), (16) can be expressed as

$$P_e\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b\right) = \mathcal{N}_2 \sum_{(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}} \Pr\left(y_j | s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}}\right) \prod_{\substack{i'=1 \\ i'\neq i}}^{2m} P_a\left(\left[\mathbf{v}^{\mathrm{AB}}\left(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}}\right)\right]_{i'}\right) \tag{17}$$

The extrinsic probabilities $P_e(\mathbf{v}_{j,i}^{\mathrm{AB}})$ will then be deinterleaved and parallel-to-serial (P/S) converted. They will be mapped to the *a priori* probabilities of $c_{t'}^{\xi}$ by

$$P_e\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b\right) \mapsto P_a\left(c_{t'}^{\xi} = b\right),\ \xi \in \{\mathrm{A}, \mathrm{B}\}. \tag{18}$$

Based on the joint trellis, the MAP decoding algorithm is performed to determine the APP of the end nodes' coded bits $P_p(c_{t'}^{\xi} = b|\mathbf{y})$ and information bits $P_p(u_t^{\xi} = b|\mathbf{y})$.

In the following iterations, the extrinsic probability of $c_{t'}^{\xi}$ can be determined by

$$P_e\left(c_{t'}^{\xi} = b\right) = \mathcal{N}_3 \frac{P_p\left(c_{t'}^{\xi} = b|\mathbf{y}\right)}{P_a\left(c_{t'}^{\xi} = b\right)}, \tag{19}$$

where $\mathcal{N}_3 = (\sum_{b\in\{0,1\}} P_e(c_{t'}^{\xi} = b))^{-1}$ is again a normalization factor. They will then be S/P converted and interleaved, and further mapped to the *a priori* probabilities of $\mathbf{v}_{j,i}^{\mathrm{AB}}$ as

$$P_e\left(c_{t'}^{\xi} = b\right) \mapsto P_a\left(\mathbf{v}_{j,i}^{\mathrm{AB}} = b\right),\ \xi \in \{\mathrm{A}, \mathrm{B}\}. \tag{20}$$

With the newly updated probabilities $P_a(\mathbf{v}_{j,i}^{\mathrm{AB}})$, another iteration starts with the demapper re-calculating the extrinsic probabilities $P_e(\mathbf{v}_{j,i}^{\mathrm{AB}})$ as in (17). After a certain number of iterations, the iterative system terminates and the information bits u_t^{ξ} can be estimated by

$$\begin{cases} u_t^{\xi} = 1, & \text{if } P_p\left(u_t^{\xi} = 1|\mathbf{y}\right) > P_p\left(u_t^{\xi} = 0|\mathbf{y}\right), \\ u_t^{\xi} = 0, & \text{if } P_p\left(u_t^{\xi} = 1|\mathbf{y}\right) < P_p\left(u_t^{\xi} = 0|\mathbf{y}\right). \end{cases} \tag{21}$$

Finally, R determines the message sequence $\mathbf{u}_{\mathrm{AB}}$ and generates the NC message $\mathbf{u}_N$ for the BC phase transmission.

Note that the proposal is a general iterative demapping-decoding approach for TWR channels with PNC. Its applications are not limited to QAM modulations. Asymmetric setup where the end nodes employ modulations of different order can also be accommodated. For example, node A may employ 8PSK while node B employs 16QAM. In this case, the set of interleaved coded bits is $\mathbf{v}_j^{\mathrm{AB}} = [v_{j,1}^{\mathrm{A}}, v_{j,2}^{\mathrm{A}}, v_{j,3}^{\mathrm{A}}, v_{j,1}^{\mathrm{B}}, v_{j,2}^{\mathrm{B}}, v_{j,3}^{\mathrm{B}}, v_{j,4}^{\mathrm{B}}]$. Transmitted symbol set $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}$ will be redefined according to the 8PSK and 16QAM constellations. Based on Section III-D, we can see the demapper requires individual mapping constellation of the end nodes. During the demapping, one transmitted symbol is seen as an interference to the other. Therefore, it is permissible that $\log_2|\chi_{\mathrm{A}}| \neq \log_2|\chi_{\mathrm{B}}|$.

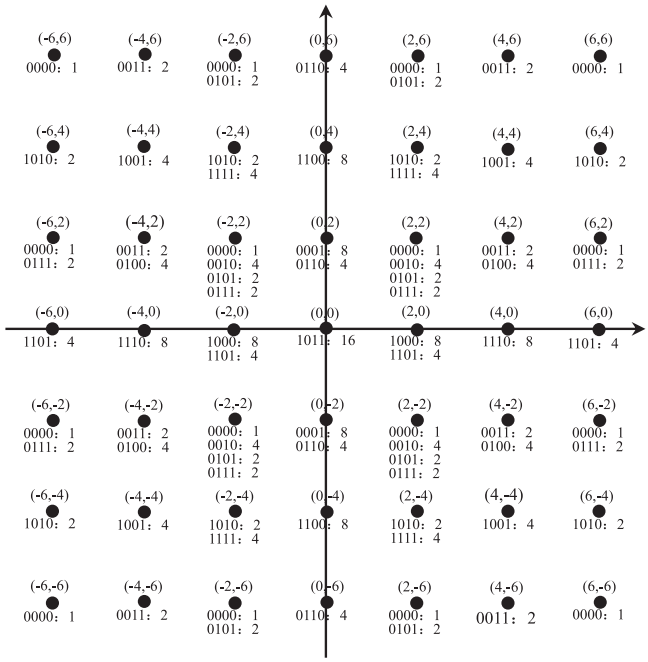

Fig. 5. Superimposed constellation of two MSEW 16QAMs.

D. Resolving the Demapping Ambiguity

It is important to point out that over the TWR Gaussian channel where $h_j^{\mathrm{A}} = h_j^{\mathrm{B}} = 1, \forall j$, demapping ambiguity often occurs especially when high order modulation schemes are employed by the end nodes. In the absence of fading coefficients, one superimposed symbol may correspond to multiple transmitted symbol pairs. They carry different sets of interleaved coded bits, causing the demapping ambiguity. In particular, if both end nodes employ M-QAMs, the relay may obtain less than M^2 distinct channel observations $\Pr(y_j|s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$. Fig. 5 shows the superimposed signal constellation resulted from two MSEW 16QAMs whose constellation is shown in Fig. 3. There are only 49 distinct superimposed symbols $s_\varrho^{\mathrm{A}} + s_\sigma^{\mathrm{B}}$. For example, symbol $(-2, 6)$ can be the superposition outcome between s_1^{A} and $s_1^{\mathrm{B}}, s_2^{\mathrm{A}}$ and s_7^{B}, or s_7^{A} and s_2^{B}. Those symbol pairs carry interleaved coded bits of [0 0 0 1 0 0 0 1], [0 0 1 0 0 1 1 1], and [0 1 1 1 0 0 1 0] respectively. Without knowledge of the transmitted symbol pair probability $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$, the APPs $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}}|y_j)$ will be the same for those symbol pairs, resulting in demapping ambiguity. The proposed BICM-ID scheme resolves this ambiguity by providing the decoding feedback which form the transmitted symbol pair probabilities as in (15).

E. Complexity

In the above iterative scheme, decoding estimates the end nodes' coded bits. It leads to a more accurate estimation of the transmitted symbol pair probability $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$ that is iteratively updated, during which it is crucial to determine $P_e(\mathbf{v}_{j,i}^{\mathrm{AB}} = b)$ of (17). For this, the demapper requires the end nodes' mapping constellations in order to identify the transmitted symbol pairs $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}$. For example, if both end nodes employ the MSEW 16QAM of Fig. 3, to determine $P_e(\mathbf{v}_{j,5}^{\mathrm{AB}} = 1)$ (or $P_e(v_{j,1}^{\mathrm{B}} = 1)$), the demapper needs to identify the constellation points of set $\chi_{1,1}^{\mathrm{B}}$. Let $s_\sigma^{\mathrm{B}} \in \chi_{1,1}^{\mathrm{B}}$, each identified transmitted symbol s_σ^{B} can be paired with 16 points of χ_{A}, constituting 16 transmitted symbol pairs $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{5,1}$. Since $|\chi_{1,1}^{\mathrm{B}}| = 8$, there are 128 symbol pairs in $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{5,1}$. In general, in order to determine a particular $P_e(\mathbf{v}_{j,i}^{\mathrm{AB}} = b)$, we have to identify $\frac{M}{2} \cdot M = \frac{M^2}{2}$ transmitted symbol pairs of $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}$. Based on (14), we know that updating each probability $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$ requires $2m$ multiplications. Hence, the complexity in computing each $P_e(\mathbf{v}_{j,i}^{\mathrm{AB}} = b)$ will be $O(mM^2)$. Considering the length of codeword sequence $\mathbf{c}_{\mathrm{AB}}$, the overall demapping complexity will be scaled up by a factor of $\frac{2l}{r}$.

IV. Mapping Constellation Design

This section proposes the end nodes' constellation design criterion, aiming to optimize the BICM-ID performance. The following analysis is performed under the content of TWR fast fading channel in which each transmitted symbol experiences an independent fading.

The average decoding error probability can be expressed as [30]

$$P_b \le \frac{1}{2l} \sum_{q=1}^{2l} W(q) \Pr(\mathbf{c}_{\mathrm{AB}} \to \mathbf{e}_{\mathrm{AB}}), \tag{22}$$

where $\Pr(\mathbf{c}_{\mathrm{AB}} \to \mathbf{e}_{\mathrm{AB}})$ is the pairwise error probability (PEP) of the end nodes' coded bit sequence $\mathbf{c}_{\mathrm{AB}}$, and

$$\begin{aligned}\mathbf{e}_{\mathrm{AB}} = \big[&e_1^{\mathrm{A}}, e_2^{\mathrm{A}}, \ldots, e_n^{\mathrm{A}}, e_1^{\mathrm{B}}, e_2^{\mathrm{B}}, \ldots, e_n^{\mathrm{B}}, \ldots\ldots, \\ &e_{\frac{l}{r}-n+1}^{\mathrm{A}}, e_{\frac{l}{r}-n+2}^{\mathrm{A}}, \ldots, e_{\frac{l}{r}}^{\mathrm{A}}, e_{\frac{l}{r}-n+1}^{\mathrm{B}}, e_{\frac{l}{r}-n+2}^{\mathrm{B}}, \ldots, e_{\frac{l}{r}}^{\mathrm{B}}\big]\end{aligned} \tag{23}$$

denotes its erroneous estimation. Moreover, q is the Hamming distance between $\mathbf{c}_{\mathrm{AB}}$ and $\mathbf{e}_{\mathrm{AB}}$, and $W(q)$ is the weight of those error events. With an ideal random interleaver and the ML decoding algorithm, $\Pr(\mathbf{c}_{\mathrm{AB}} \to \mathbf{e}_{\mathrm{AB}})$ can be determined by [30]

$$\Pr(\mathbf{c}_{\mathrm{AB}} \to \mathbf{e}_{\mathrm{AB}}) = \left(\Pr\left(c_{t'}^{\xi} \to e_{t'}^{\xi}\right)\right)^q, \ \xi \in \{\mathrm{A}, \mathrm{B}\} \tag{24}$$

where $\Pr(c_{t'}^{\xi} \to e_{t'}^{\xi})$ is the PEP between $c_{t'}^{\xi}$ and its erroneous estimation $e_{t'}^{\xi}$. Integrating this paper's notation fashion, it can be written as

$$\Pr\left(c_{t'}^{\xi} \to e_{t'}^{\xi}\right) = \Pr\left(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}}\right), \ \xi \in \{\mathrm{A}, \mathrm{B}\} \tag{25}$$

where $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}})$ is the PEP between $\mathbf{v}_{j,i}^{\mathrm{AB}}$ and its erroneous version $\mathbf{e}_{j,i}^{\mathrm{AB}}$.

Generating the XOR NC message leads to the average NC message bit error probability of

$$P_N = (1 - P_b)P_b + P_b(1 - P_b) = 2P_b(1 - P_b). \tag{26}$$

Hence, in order to optimize the decoding performance, one should minimize P_b. By substituting (24) and (25) into (22), we obtain

$$P_b \le \frac{1}{2l} \sum_{q=1}^{2l} W(q) \left(\Pr\left(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}}\right)\right)^q. \tag{27}$$

Therefore, the following constellation design criterion is derived in order to minimize $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}})$.

Let us denote $(x_j^{\mathrm{A}}, x_j^{\mathrm{B}})$ as a particular transmitted symbol pair of $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}$. In BICM-ID, if $(x_j^{\mathrm{A}}, x_j^{\mathrm{B}})$ is erroneously estimated as another symbol pair $(w_j^{\mathrm{A}}, w_j^{\mathrm{B}})$ of $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,\bar{b}}$, where $[\Lambda_{\mathrm{A}}^{-1}(w_j^{\mathrm{A}}), \Lambda_{\mathrm{B}}^{-1}(w_j^{B})]_{i'} = [\Lambda_{\mathrm{A}}^{-1}(x_j^{\mathrm{A}}), \Lambda_{\mathrm{B}}^{-1}(x_j^{\mathrm{B}})]_{i'}$ for $i' = 1, 2, \ldots, 2m$ and $i' \neq i$, the demapper will make an estimation error for bit $\mathbf{v}_{j,i}^{\mathrm{AB}}$. Let $\Pr((x_j^{\mathrm{A}}, x_j^{\mathrm{B}}) \to (w_j^{\mathrm{A}}, w_j^{\mathrm{B}}))$ denote the PEP of the transmitted symbol pairs, $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}})$ can be determined by

$$\begin{aligned}&\Pr\left(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}}\right) \\ &\quad= \frac{1}{2m4^m} \sum_{i=1}^{2m} \sum_{b=0}^{1} \sum_{\substack{(x_j^{\mathrm{A}}, x_j^{\mathrm{B}}) \in \\ (s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}}} \Pr\left(\left(x_j^{\mathrm{A}}, x_j^{\mathrm{B}}\right) \to \left(w_j^{\mathrm{A}}, w_j^{\mathrm{B}}\right)\right).\end{aligned} \tag{28}$$

In general, $(w_j^{\mathrm{A}}, w_j^{\mathrm{B}})$ is not unique in $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,\bar{b}}$. $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}})$ can be further expressed as

$$\begin{aligned}&\Pr\left(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}}\right) \\ &\le \frac{1}{2m4^m} \sum_{i=1}^{2m} \sum_{b=0}^{1} \sum_{\substack{(w_j^{\mathrm{A}}, w_j^{\mathrm{B}}) \in \\ (s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,\bar{b}}}} \sum_{\substack{(x_j^{\mathrm{A}}, x_j^{\mathrm{B}}) \in \\ (s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,b}}} \Pr\left(\left(x_j^{\mathrm{A}}, x_j^{\mathrm{B}}\right) \to \left(w_j^{\mathrm{A}}, w_j^{\mathrm{B}}\right)\right),\end{aligned} \tag{29}$$

where the equality holds when $(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})_{i,\bar{b}} = \{(w_j^{\mathrm{A}}, w_j^{\mathrm{B}})\}$.

Since the following analysis dose not need to specify a particular transmitted symbol pair, we will drop symbol index j for the sake of simplicity.

Based on (11), we know that the demapper obtains channel observations by the distance metric

$$\mathrm{D}\left(y, \left(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}}\right)\right) = \|y - h^{\mathrm{A}} s_\varrho^{\mathrm{A}} - h^{\mathrm{B}} s_\sigma^{\mathrm{B}}\|^2. \tag{30}$$

The demapper will favor symbol pair $(w^{\mathrm{A}}, w^{\mathrm{B}})$ over $(x^{\mathrm{A}}, x^{\mathrm{B}})$ if $\mathrm{D}(y, (w^{\mathrm{A}}, w^{\mathrm{B}})) < \mathrm{D}(y, (x^{\mathrm{A}}, x^{\mathrm{B}}))$. Therefore, $\Pr((x^{\mathrm{A}}, x^{\mathrm{B}}) \to (w^{\mathrm{A}}, w^{\mathrm{B}}))$ can be reinterpreted by

$$\begin{aligned}&\Pr\left(\left(x^{\mathrm{A}}, x^{\mathrm{B}}\right) \to \left(w^{\mathrm{A}}, w^{\mathrm{B}}\right)\right) \\ &\quad= \Pr\left(\mathrm{D}\left(y, \left(w^{\mathrm{A}}, w^{\mathrm{B}}\right)\right) < \mathrm{D}\left(y, \left(x^{\mathrm{A}}, x^{\mathrm{B}}\right)\right)\right).\end{aligned} \tag{31}$$

With the complex signal model, we write $h^{\xi} = h_1^{\xi} + \mathrm{j}h_2^{\xi}$, $x^{\xi} = x_1^{\xi} + \mathrm{j}x_2^{\xi}$, $w^{\xi} = w_1^{\xi} + \mathrm{j}w_2^{\xi}$ and $z = z_1 + \mathrm{j}z_2$, where

$\mathrm{j}=\sqrt{-1}$. The discrepancy between $\mathrm{D}(y,(w^{\mathrm{A}},w^{\mathrm{B}}))$ and $\mathrm{D}(y,(x^{\mathrm{A}},x^{\mathrm{B}}))$ will be

$$\begin{aligned}\Delta &= \mathrm{D}\left(y,\left(w^{\mathrm{A}},w^{\mathrm{B}}\right)\right)-\mathrm{D}\left(y,\left(x^{\mathrm{A}},x^{\mathrm{B}}\right)\right)\\ &= \|h^{\mathrm{A}}\left(x^{\mathrm{A}}-w^{\mathrm{A}}\right)+h^{\mathrm{B}}\left(x^{\mathrm{B}}-w^{\mathrm{B}}\right)\|^2+2\left(h_1^{\mathrm{A}}\left(x_1^{\mathrm{A}}-w_1^{\mathrm{A}}\right)\right.\\ &\quad\left.+h_1^{\mathrm{B}}\left(x_1^{\mathrm{B}}-w_1^{\mathrm{B}}\right)-h_2^{\mathrm{A}}\left(x_2^{\mathrm{A}}-w_2^{\mathrm{A}}\right)-h_2^{\mathrm{B}}\left(x_2^{\mathrm{B}}-w_2^{\mathrm{B}}\right)\right)z_1\\ &\quad+2\left(h_1^{\mathrm{A}}\left(x_2^{\mathrm{A}}-w_2^{\mathrm{A}}\right)+h_1^{\mathrm{B}}\left(x_2^{\mathrm{B}}-w_2^{\mathrm{B}}\right)+h_2^{\mathrm{A}}\left(x_1^{\mathrm{A}}-w_1^{\mathrm{A}}\right)\right.\\ &\quad\left.+h_2^{\mathrm{B}}\left(x_1^{\mathrm{B}}-w_1^{\mathrm{B}}\right)\right)z_2.\end{aligned} \tag{32}$$

By defining

$$d_\xi = x^\xi - w^\xi, \tag{33}$$

$\Pr((x^{\mathrm{A}},x^{\mathrm{B}})\to(w^{\mathrm{A}},w^{\mathrm{B}}))$ of (31) can be expressed as

$$\begin{aligned}&\Pr\left(\left(x^{\mathrm{A}},x^{\mathrm{B}}\right)\to\left(w^{\mathrm{A}},w^{\mathrm{B}}\right)\right)=\Pr\left(\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{2}\right.\\ &<\left(h_1^{\mathrm{A}}\left(w_1^{\mathrm{A}}-x_1^{\mathrm{A}}\right)+h_1^{\mathrm{B}}\left(w_1^{\mathrm{B}}-x_1^{\mathrm{B}}\right)-h_2^{\mathrm{A}}\left(w_2^{\mathrm{A}}-x_2^{\mathrm{A}}\right)\right.\\ &\left.-h_2^{\mathrm{B}}\left(w_2^{\mathrm{B}}-x_2^{\mathrm{B}}\right)\right)z_1+\left(h_1^{\mathrm{A}}\left(w_2^{\mathrm{A}}-x_2^{\mathrm{A}}\right)+h_1^{\mathrm{B}}\left(w_2^{\mathrm{B}}-x_2^{\mathrm{B}}\right)\right.\\ &\left.\left.+h_2^{\mathrm{A}}\left(w_1^{\mathrm{A}}-x_1^{\mathrm{A}}\right)+h_2^{\mathrm{B}}\left(w_1^{\mathrm{B}}-x_1^{\mathrm{B}}\right)\right)z_2\right).\end{aligned} \tag{34}$$

Since z_1 and z_2 are Gaussian random variables with variance $\frac{N_0}{2}$, $(h_1^{\mathrm{A}}(w_1^{\mathrm{A}}-x_1^{\mathrm{A}})+h_1^{\mathrm{B}}(w_1^{\mathrm{B}}-x_1^{\mathrm{B}})-h_2^{\mathrm{A}}(w_2^{\mathrm{A}}-x_2^{\mathrm{A}})-h_2^{\mathrm{B}}(w_2^{\mathrm{B}}-x_2^{\mathrm{B}}))z_1+(h_1^{\mathrm{A}}(w_2^{\mathrm{A}}-x_2^{\mathrm{A}})+h_1^{\mathrm{B}}(w_2^{\mathrm{B}}-x_2^{\mathrm{B}})+h_2^{\mathrm{A}}(w_1^{\mathrm{A}}-x_1^{\mathrm{A}})+h_2^{\mathrm{B}}(w_1^{\mathrm{B}}-x_1^{\mathrm{B}}))z_2$ is also a Gaussian random variable with variance $\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2N_0}{2}$.

Let $h^\xi=\alpha^\xi e^{\mathrm{j}\theta_\xi}$, where α^ξ is the fading amplitude that follows Rayleigh distribution with pdf $f(\alpha^\xi)=2\alpha^\xi e^{-(\alpha^\xi)^2}$ and phase $\theta_\xi\in[0,2\pi)$. Using *Lemma 3* that is stated in Appendix A, we know

$$\begin{aligned}&\Pr\left(\left(x^{\mathrm{A}},x^{\mathrm{B}}\right)\to\left(w^{\mathrm{A}},w^{\mathrm{B}}\right)\right)\\ &=\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{Q\left(\frac{\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{2}}{\sqrt{\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2N_0}{2}}}\right)\right\}\\ &<\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\},\end{aligned} \tag{35}$$

where

$$\begin{aligned}&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}\\ &=\int_0^{+\infty}\int_0^{+\infty}\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)f\left(a^{\mathrm{A}}\right)f\left(a^{\mathrm{B}}\right)da^{\mathrm{A}}da^{\mathrm{B}}.\end{aligned} \tag{36}$$

Now, we will derive a closed form expression for the upper bound of (36). Since d_{A} and d_{B} are complex variables that can be written as $d_{\mathrm{A}}=d_1^{\mathrm{A}}+\mathrm{j}d_2^{\mathrm{A}}$ and $d_{\mathrm{B}}=d_1^{\mathrm{B}}+\mathrm{j}d_2^{\mathrm{B}}$, respectively, $\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2$ can be manipulated into

$$\begin{aligned}&\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2\\ &=\left(a^{\mathrm{A}}\right)^2\|d_{\mathrm{A}}\|^2+2\Gamma\left(d_{\mathrm{A}},d_{\mathrm{B}}\right)a^{\mathrm{A}}a^{\mathrm{B}}+\left(a^{\mathrm{B}}\right)^2\|d_{\mathrm{B}}\|^2,\end{aligned} \tag{37}$$

where

$$\Gamma\left(d_{\mathrm{A}},d_{\mathrm{B}}\right)=d_1^{\mathrm{A}}d_1^{\mathrm{B}}+d_2^{\mathrm{A}}d_2^{\mathrm{B}}. \tag{38}$$

Further let

$$\Upsilon_\xi=4N_0+\|d_\xi\|^2, \tag{39}$$

we can introduce the following Lemma that characterizes the upper bound of (36).

Lemma 1: If $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})\geq 0$,

$$\begin{aligned}&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}\\ &<\frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}-\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}.\end{aligned} \tag{40}$$

If $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})<0$,

$$\begin{aligned}&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}<\\ &\frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}-\frac{16\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}+\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}.\end{aligned} \tag{41}$$

Proof: Please see Appendix B. ■

Further recalling (29), we can conclude that $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}}\to\mathbf{e}_{j,i}^{\mathrm{AB}})$ is upper bounded by

$$\begin{aligned}&\Pr\left(\mathbf{v}_{j,i}^{\mathrm{AB}}\to\mathbf{e}_{j,i}^{\mathrm{AB}}\right)\\ &<\frac{1}{2m4^m}\sum_{i=1}^{2m}\sum_{b=0}^{1}\sum_{\substack{(x^{\mathrm{A}},x^{\mathrm{B}})\in\\(s_\varrho^{\mathrm{A}},s_\sigma^{\mathrm{B}})_{i,b}}}\sum_{\substack{(w^{\mathrm{A}},w^{\mathrm{B}})\in\\(s_\varrho^{\mathrm{A}},s_\sigma^{\mathrm{B}})_{i,\bar{b}}}}\mathcal{D}(d_{\mathrm{A}},d_{\mathrm{B}}),\end{aligned} \tag{42}$$

where

$$\mathcal{D}(d_{\mathrm{A}},d_{\mathrm{B}})=\begin{cases}\frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}-\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}, & \text{if } \Gamma(d_{\mathrm{A}},d_{\mathrm{B}})\geq 0;\\ \frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}-\frac{16\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}} & \\ \quad+\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}, & \text{if } \Gamma(d_{\mathrm{A}},d_{\mathrm{B}})<0.\end{cases} \tag{43}$$

It can be seen that $\mathcal{D}(d_{\mathrm{A}},d_{\mathrm{B}})$ is a function of d_ξ, Υ_ξ and SNR, where d_ξ and Υ_ξ are determined by the constellation χ_ξ of the end nodes. Therefore, we can optimize the BICM-ID performance by designing the constellations χ_{A} and χ_{B}. The following Theorem presents the constellation design criterion for BICM-ID over the TWR fading channel.

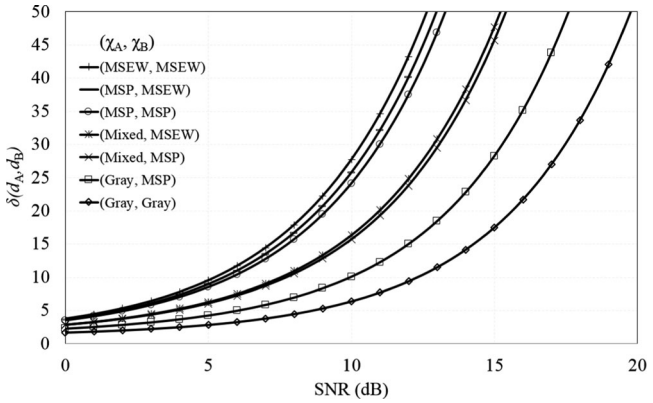

Fig. 6. $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ for different combinations of 16QAM constellations.

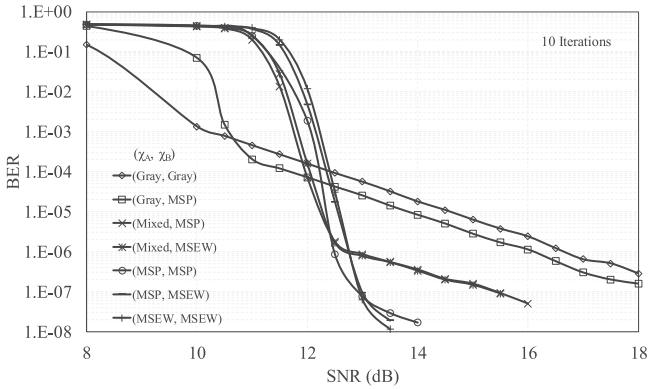

Fig. 7. BICM-ID performance for different combinations of 16QAM constellations.

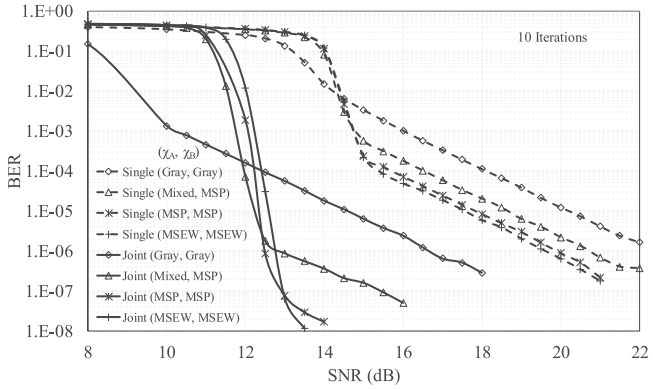

Fig. 8. Performance comparison between using the joint trellis and the single trellis.

Theorem 2: In order to optimize the BICM-ID performance, constellations χ_{A} and χ_{B} should be designed such that

$$\delta(d_{\mathrm{A}}, d_{\mathrm{B}}) = \left(\frac{1}{2m4^m} \sum_{i=1}^{2m} \sum_{b=0}^{1} \sum_{\substack{(x^{\mathrm{A}},x^{\mathrm{B}})\in \\ (s_\varrho^{\mathrm{A}},s_\sigma^{\mathrm{B}})_{i,b}}} \sum_{\substack{(w^{\mathrm{A}},w^{\mathrm{B}})\in \\ (s_\varrho^{\mathrm{A}},s_\sigma^{\mathrm{B}})_{i,\bar{b}}}} \mathcal{D}(d_{\mathrm{A}}, d_{\mathrm{B}}) \right)^{-1} \tag{44}$$

is maximized.

Proof: Based on (42), we can see that by maximizing $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$, the upper bound of $\Pr(\mathbf{v}_{j,i}^{\mathrm{AB}} \to \mathbf{e}_{j,i}^{\mathrm{AB}})$ can be minimized. Consequently, BICM-ID performance will be optimized. ■

Based on the above analysis, the design metric $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ is the reciprocal of the upper bound for the PEP of interleaved coded bits $\mathbf{v}_{j,i}^{\mathrm{AB}}$. Therefore, in order to optimize the BICM-ID performance, we should choose constellations χ_{A} and χ_{B} such that they can maximize $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ over the interested SNR region. However, if both of the end nodes employ 16QAM, there are $(16!)^2 \cong 4.38 \times 10^{26}$ possible designs of $(\chi_{\mathrm{A}}, \chi_{\mathrm{B}})$. It is infeasible to deploy a brutal search for the optimal combination of $(\chi_{\mathrm{A}}, \chi_{\mathrm{B}})$. In this paper, we consider the known 16QAM constellations, including the Gray mapping, the mixed mapping [25], the modified set-partitioning (MSP) mapping [25] and the MSEW mapping [26].

Fig. 6 shows the numerical results of $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ over the SNR region of 0–20 dB for several 16QAM combinations. It can be seen that if both of the end nodes employ the MSEW 16QAM, its $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ prevails the others in the SNR region. In contrast, the (Gray, Gray) combination will be undesirable. This is in concert with the existing conclusion on BICM-ID over the point-to-point channel [31]. Intuitively, over the TWR channels one user's transmitted symbol is considered as an interference to the other's. The proposed BICM-ID estimates both end nodes' information instead of their network coded information. This is analog to the point-to-point channel where one user's signal is corrupted by user interference and noise. Therefore, a constellation that is good for the point-to-point channel will also be good for the TWR channels. Fig. 7 further shows the bit error rate (BER) performance of the BICM-ID scheme that employs the above shown 16QAM combinations. The $(5, 7)_8$ convolutional code is used and the BICM-ID is performed with ten iterations. It shows the asymptotic performances agree with the results of Fig. 6. Among the seven combinations, the (MSEW, MSEW) yields the best asymptotic performance. Note that our design criterion given in Theorem 2 is concluded under the assumption of perfect decoding feedback. Effectiveness of the $\delta(d_{\mathrm{A}}, d_{\mathrm{B}})$ prediction is realized in BER curve's asymptotic region.

To show the impact of using a joint trellis for the MAP decoding, Fig. 8 compares the BICM-ID performance between using a joint trellis and a single trellis. Again, it is assumed that both end nodes employ the $(5, 7)_8$ convolutional code. It shows that the joint trellis based decoding obtains substantial performance improvement and yields an earlier decoding convergence and lower error floor. This is because the joint trellis based decoding estimates both of the end nodes' coded bits, improving estimation of the transmitted symbol pair probability $\Pr(s_\varrho^{\mathrm{A}}, s_\sigma^{\mathrm{B}})$.

V. EXIT Analysis

This section utilizes the EXIT chart [24] to analyze the convergence behavior of the proposed iterative scheme. Let T_1 and T_2 denote the *a priori*-extrinsic transfer functions of the demapper and the MAP decoder, respectively. Further let $I_{a_1}^{\xi}$ and $I_{e_1}^{\xi}$ denote the *a priori* and extrinsic mutual information of node ξ, which are observed before and after the demapping,

respectively. Similarly, $I_{a_2}^{\mathrm{AB}}$ and $I_{e_2}^{\mathrm{AB}}$ denote the *a priori* and extrinsic mutual information, which are observed before and after the MAP decoding, respectively. Note that unlike the demapper, the decoder treats the end nodes' codeword as an entity $\mathbf{c}_{\mathrm{AB}}$ with which we do not need to distinguish the mutual information between the two nodes.

For the demapper, $I_{e_1}^{\xi}$ can be seen as a function of $I_{a_1}^{\mathrm{A}}$, $I_{a_1}^{\mathrm{B}}$ and $\frac{E_b}{N_0}$, i.e.,

$$I_{e_1}^{\xi} = T_1\left(I_{a_1}^{\mathrm{A}}, I_{a_1}^{\mathrm{B}}, \frac{E_b}{N_0}\right). \tag{45}$$

After de-interleaving and the P/S conversion, $I_{e_1}^{\mathrm{A}}$ and $I_{e_1}^{\mathrm{B}}$ are mapped to $I_{a_2}^{\mathrm{AB}}$. Subsequently, $I_{e_2}^{\mathrm{AB}}$ can be generated by

$$I_{e_2}^{\mathrm{AB}} = T_2\left(I_{a_2}^{\mathrm{AB}}\right). \tag{46}$$

After the S/P conversion and interleaving, $I_{e_2}^{\mathrm{AB}}$ is mapped back to $I_{a_1}^{\mathrm{A}}$ and $I_{a_1}^{\mathrm{B}}$. During these *a priori*-extrinsic transfers, mutual information $I_{e_1}^{\xi}$ is determined through simulation by

$$I_{e_1}^{\xi} = 1 - \frac{r}{l}\sum_{j=1}^{\frac{l}{mr}}\sum_{i=1}^{m}\mathcal{H}_b(v_{j,i}^{\xi}), \tag{47}$$

where $\mathcal{H}_b(v_{j,i}^{\xi})$ is the binary entropy function

$$\mathcal{H}_b\left(v_{j,i}^{\xi}\right) = -\sum_{b=0}^{1} P_e\left(v_{j,i}^{\xi} = b\right)\log_2\left(P_e\left(v_{j,i}^{\xi} = b\right)\right). \tag{48}$$

Similarly, mutual information $I_{e_2}^{\mathrm{AB}}$ is determined by

$$I_{e_2}^{\mathrm{AB}} = 1 - \frac{r}{2l}\sum_{t=1}^{\frac{2l}{r}}\mathcal{H}_b\left(c_t^{\mathrm{AB}}\right), \tag{49}$$

where $\mathcal{H}_b(c_t^{\mathrm{AB}})$ is calculated as in (48) by replacing $P_e(v_{j,i}^{\xi} = b)$ with $P_e(c_t^{\mathrm{AB}} = b)$.

Fig. 9 shows the EXIT chart with both nodes employing the $(5, 7)_8$ convolutional code and the MSEW 16QAM over the TWR fading channel. It shows when SNR = 10.8 dB, EXIT curve of the demapper is tangent with the inverse EXIT curve of the decoder. The iterative system starts to function in message recovery and its BER curve's waterfall will begin to appear. This SNR is often marked as the pinch off SNR limit denoted $\mathrm{SNR}_{\mathrm{off}}$. By increasing the SNR, the tunnel between the two curves widens, implying an enhanced convergence behavior of the iterative scheme. The trajectories show less iterations will be needed by increasing the SNR.

The EXIT chart can be further utilized to choose channel codes that can produce a better iterative convergence. For example, Table I shows the $\mathrm{SNR}_{\mathrm{off}}$ for the BICM-ID scheme with different convolutional codes and 16QAM constellations, in which both end nodes employ the same code. It shows with the same 16QAM combination, the $(5, 7)_8$ code yields a better convergence than the $(13, 15)_8$ code. On the other hand, with the same code, the (Gray, MSP) combination yields a better convergence than the (MSEW, MSEW) combination. Fig. 10 further validates Table I's results by showing the BER performances. Despite the $(5, 7)_8$ code can yield a better convergence, the $(13, 15)_8$ code can however produce a better asymptotic performance thanks to its better distance property. Echoing our analysis of Section IV, the (MSEW, MSEW) combination produces a better asymptotic performance despite its convergence disadvantage. In summary, the constellation design analysis guides the choice of constellations in maximizing the asymptotic performance, while the EXIT analysis guilds the choices of channel codes and constellations in optimizing the iterative convergence.

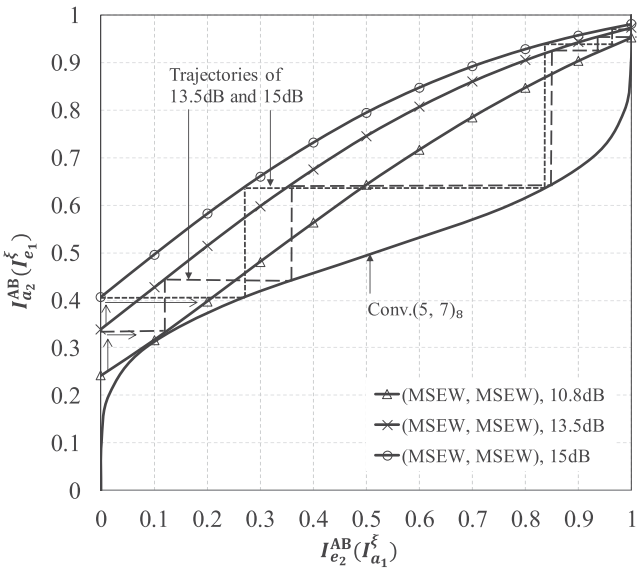

Fig. 9. EXIT chart of the proposed BICM-ID over the TWR fading channel.

TABLE I
$\mathrm{SNR}_{\mathrm{OFF}}$ OF DIFFERENT CONVOLUTIONAL CODES AND 16QAM CONSTELLATIONS

Convolutional codes	16QAM constellations	$\mathrm{SNR}_{\mathrm{off}}$
$(5, 7)_8$	(Gray, MSP)	8.8 dB
	(MSEW, MSEW)	10.8 dB
$(13, 15)_8$	(Gray, MSP)	9.7 dB
	(MSEW, MSEW)	11.5 dB

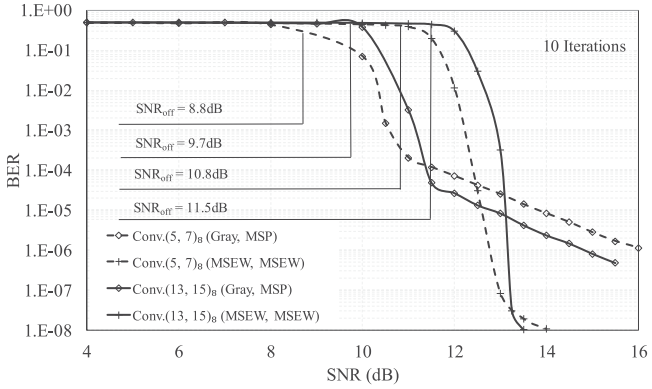

Fig. 10. BICM-ID performance of different codes and 16QAM combinations.

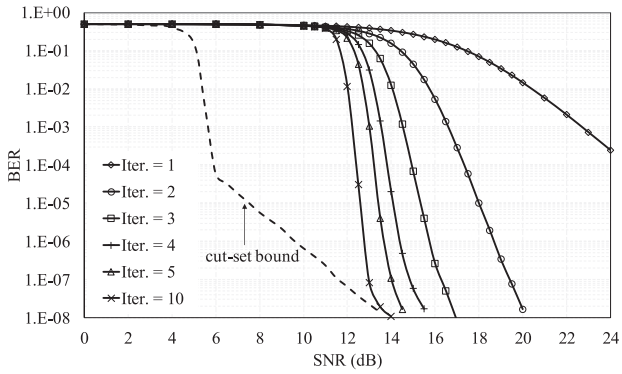

Fig. 11. BICM-ID performance over the TWR fading channel with $(5, 7)_8$ convolutional code and (MSEW, MSEW) 16QAM combination.

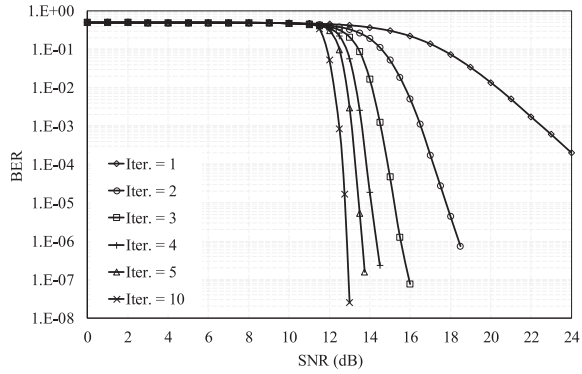

Fig. 12. BICM-ID performance over the TWR fading channel with $(5, 7)_8$ and $(13, 15)_8$ convolutional codes and (MSP, MSEW) 16QAM combination.

VI. Simulation Results

This section further supplements other simulation results of BICM-ID over the TWR fast fading channel. Fig. 11 shows the BER performance of BICM-ID in which both end nodes employ the $(5, 7)_8$ convolutional code and the MSEW 16QAM. It shows decoding performance can be significantly improved by increasing the number of iterations. The feedbacks of the MAP decoder help produce a more accurate estimation on the transmitted symbol pair probability. It enables the demapper produce a better estimation on the end nodes' coded bits, which in return helps strengthen the decoding performance. The cut-set performance limit is also shown as a benchmark. It is obtained by assuming the end nodes transmit in orthogonal time slots during the MAC phase. Consequently, relay will decode node ξ's message based on its received signal $y_j^{\xi} = h_j^{\xi} x_j^{\xi} + z_j$. Although the cut-set performance and the PNC based decoding performance converge asymptotically, the BICM-ID performance still exhibits a few dB gap to the cut-set limit in the waterfall region. This is caused by the PNC operation that superimposes two transmitted symbols. During the demapping, one becomes an interference to the other. When the SNR is not high enough, channel coding cannot help dissolve this interference. Narrowing the performance gap to the cut-set limit continues to sustain our endeavor in which the capacity-approaching codes may be considered. It is important to point out that the proposed BICM-ID does not impose coherent coding and modulation schemes for the end nodes. In order to substantiate this advantage, Fig. 12 shows the BICM-ID performance where node A employs the $(5, 7)_8$ convolutional code and the MSP 16QAM while node B employs the $(13, 15)_8$ code and the MSEW 16QAM. Again, it shows significant iterative decoding gain can be achieved. It also outperforms the results of Fig. 11 where the $(5, 7)_8$ code is employed by both of the end nodes. Based on the analysis of Section IV, we understand that the (MSP, MSEW) and (MSEW, MSEW) yield similar BICM-ID performance. Hence, this improved margin is mainly due to a stronger channel code is employed by node B.

VII. Conclusion

This paper has proposed a general iterative demapping-decoding approach, namely BICM-ID, for the TWR channels with PNC. It realizes spectrally efficient message exchange between two end nodes. The joint trellis based MAP decoding estimates both end nodes' information and coded bits. It enables the demapper produce more accurate transmitted symbol pair probabilities, leading to an enhanced BICM-ID performance. The proposed BICM-ID also resolves the demapping ambiguity over the TWR Gaussian channel. In order to optimize the decoding performance, constellation design criterion for the two end nodes has been proposed. It sheds light on the asymptotic performance of the proposed scheme. The EXIT analysis has been further carried out, revealing the proposal's convergence behavior over the TWR fading channel. Both of the above analyses can be utilized to better design the BICM coded TWR communication systems. Our simulation results have validated the analyses and shown remarkable coding gains can be achieved by the BICM-ID. Their asymptotic performance can also approach the cut-set performance limit. Narrowing the gap to the cut-set limit further motivates our future endeavor which will look into the deployment of capacity-approaching codes.

Appendix A
Lemma 3

Given a random variable ζ which obeys the standard normal distribution as $\mathbb{E}[\zeta] = 0$ and $var(\zeta) = 1$, the probability that ζ is greater than variable λ can be expressed as [34]

$$\Pr(\zeta > \lambda) = Q(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{\lambda}^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt, \qquad (50)$$

and $Q(\lambda)$ is bounded by

$$\frac{1}{4}\exp\left(-\lambda^2\right) < Q(\lambda) < \exp\left(-\frac{\lambda^2}{2}\right), \quad \text{if} \quad \lambda \geq 0;$$
$$Q(\lambda) < 1 - \frac{1}{4}\exp\left(-\lambda^2\right), \quad \text{if} \quad \lambda < 0. \qquad (51)$$

Appendix B
Proof of Lemma 1

By substituting (37) into (36), we have (52). Integration I in (52) can be further manipulated as (53).

By derivations, we know that integration III equates

$$\int_0^{+\infty} -d\left[\exp\left(-\frac{\Upsilon_{\mathrm{B}}(a^{\mathrm{B}})^2+2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}a^{\mathrm{B}}}{4N_0}\right)\right]=1,$$

and integration IV equates

$$\exp\left(\frac{\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})(a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}}.$$

By substituting them into (53), integration I can be written as

$$\begin{aligned}&\frac{2N_0}{\Upsilon_{\mathrm{B}}}-\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}\cdot\exp\left(\frac{\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})(a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)\\&\qquad\cdot\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}},\end{aligned}\tag{54}$$

where

$$\begin{aligned}&\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}}=\\&\sqrt{2\pi}\sqrt{\frac{2N_0}{\Upsilon_{\mathrm{B}}}}\frac{1}{\sqrt{2\pi}\sqrt{\frac{2N_0}{\Upsilon_{\mathrm{B}}}}}\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}}.\end{aligned}\tag{55}$$

Based on *Lemma 3*, we know if $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})\geq 0$,

$$\begin{aligned}&\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}}\\&\quad>\frac{1}{4}\sqrt{\frac{4\pi N_0}{\Upsilon_{\mathrm{B}}}}\exp\left(-\frac{\left(\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}\right)^2}{2N_0/\Upsilon_{\mathrm{B}}}\right).\end{aligned}\tag{56}$$

Otherwise if $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})<0$,

$$\begin{aligned}&\int_{\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}}^{+\infty}\exp\left(-\frac{(a^{\mathrm{B}})^2}{4N_0/\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{B}}\\&\quad<\sqrt{\frac{4\pi N_0}{\Upsilon_{\mathrm{B}}}}\left(1-\frac{1}{4}\exp\left(-\frac{\left(\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}\right)^2}{2N_0/\Upsilon_{\mathrm{B}}}\right)\right).\end{aligned}\tag{57}$$

Therefore, if $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})\geq 0$, the upper bound of integration I is

$$\frac{2N_0}{\Upsilon_{\mathrm{B}}}-\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{2\Upsilon_{\mathrm{B}}}\sqrt{\frac{\pi N_0}{\Upsilon_{\mathrm{B}}}}\exp\left(-\frac{(\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right).\tag{58}$$

If $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})<0$, the upper bound of integration I is

$$\begin{aligned}&\frac{2N_0}{\Upsilon_{\mathrm{B}}}-\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}\sqrt{\frac{\pi N_0}{\Upsilon_{\mathrm{B}}}}\left(2\exp\left(\frac{(-\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)\right.\\&\quad\left.-\frac{1}{2}\exp\left(\frac{(\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)\right).\end{aligned}\tag{59}$$

Based on the above conclusions, we can further pursue to obtain the upper bound of $\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\{\exp(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0})\}$. Based on (58), we know if $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})\geq 0$, $\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\{\exp(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0})\}$ is upper bounded by

$$\begin{aligned}&\frac{8N_0}{\Upsilon_{\mathrm{B}}}\underbrace{\int_0^{+\infty}a^{\mathrm{A}}\exp\left(-\frac{\Upsilon_{\mathrm{A}}(a^{\mathrm{A}})^2}{4N_0}\right)da^{\mathrm{A}}}_{\text{integration V}}-\frac{2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{\Upsilon_{\mathrm{B}}}\sqrt{\frac{\pi N_0}{\Upsilon_{\mathrm{B}}}}\cdot\\&\underbrace{\int_0^{+\infty}(a^{\mathrm{A}})^2\exp\left(-\frac{\left(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})\right)(a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{A}}}_{\text{integration VI}},\end{aligned}\tag{60}$$

$$\begin{aligned}&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}\\&=\underbrace{4\int_0^{+\infty}a^{\mathrm{A}}\exp\left(-\frac{\Upsilon_{\mathrm{A}}(a^{\mathrm{A}})^2}{4N_0}\right)\underbrace{\left[\int_0^{+\infty}a^{\mathrm{B}}\exp\left(-\frac{\Upsilon_{\mathrm{B}}(a^{\mathrm{B}})^2+2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}a^{\mathrm{B}}}{4N_0}\right)da^{\mathrm{B}}\right]}_{\text{integration I}}da^{\mathrm{A}}}_{\text{integration II}}\\&\int_0^{+\infty}\frac{2N_0}{\Upsilon_{\mathrm{B}}}\frac{\Upsilon_{\mathrm{B}}a^{\mathrm{B}}+\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}-\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{2N_0}\exp\left(-\frac{\Upsilon_{\mathrm{B}}(a^{\mathrm{B}})^2+2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}a^{\mathrm{B}}}{4N_0}\right)da^{\mathrm{B}}\\&=\frac{2N_0}{\Upsilon_{\mathrm{B}}}\underbrace{\int_0^{+\infty}\frac{\Upsilon_{\mathrm{B}}a^{\mathrm{B}}+\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{2N_0}\exp\left(-\frac{\Upsilon_{\mathrm{B}}(a^{\mathrm{B}})^2+2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}a^{\mathrm{B}}}{4N_0}\right)da^{\mathrm{B}}}_{\text{integration III}}\\&-\frac{\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}}{\Upsilon_{\mathrm{B}}}\underbrace{\int_0^{+\infty}\exp\left(-\frac{\Upsilon_{\mathrm{B}}(a^{\mathrm{B}})^2+2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})a^{\mathrm{A}}a^{\mathrm{B}}}{4N_0}\right)da^{\mathrm{B}}}_{\text{integration IV}}\end{aligned}\tag{52, 53}$$

$$\begin{aligned}
&\frac{8N_0}{\Upsilon_{\mathrm{B}}}\underbrace{\int_0^{+\infty} a^{\mathrm{A}}\exp\left(-\frac{\Upsilon_{\mathrm{A}}(a^{\mathrm{A}})^2}{4N_0}\right)da^{\mathrm{A}}}_{\text{integration V}}\\
&+\frac{2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{\Upsilon_{\mathrm{B}}}\sqrt{\frac{\pi N_0}{\Upsilon_{\mathrm{B}}}}\underbrace{\int_0^{+\infty}(a^{\mathrm{A}})^2\exp\left(-\frac{\left(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})\right)(a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{A}}}_{\text{integration VI}}\\
&-\frac{8\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{\Upsilon_{\mathrm{B}}}\sqrt{\frac{\pi N_0}{\Upsilon_{\mathrm{B}}}}\underbrace{\int_0^{+\infty}(a^{\mathrm{A}})^2\exp\left(-\frac{\left(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})\right)(a^{\mathrm{A}})^2}{4N_0\Upsilon_{\mathrm{B}}}\right)da^{\mathrm{A}}}_{\text{integration VII}}.
\end{aligned}\tag{62}$$

where integration V equates $\frac{2N_0}{\Upsilon_{\mathrm{A}}}$ and integration VI equates

$$\frac{2N_0\Upsilon_{\mathrm{B}}}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})}\sqrt{\frac{\pi N_0\Upsilon_{\mathrm{B}}}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})}}.$$

Hence, the upper bound of $\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}$ is

$$\begin{aligned}
&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}\\
&\quad<\frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}-\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}.
\end{aligned}\tag{61}$$

Based on (59), when $\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})<0$, the upper bound of $\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}$ can be written as in (62). Since integration VII equates

$$\frac{2N_0\Upsilon_{\mathrm{B}}}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})}\sqrt{\frac{\pi N_0\Upsilon_{\mathrm{B}}}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}})}},$$

together with integrations V and VI, the upper bound of $\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}$ is

$$\begin{aligned}
&\mathbb{E}_{a^{\mathrm{A}},a^{\mathrm{B}}}\left\{\exp\left(-\frac{\|h^{\mathrm{A}}d_{\mathrm{A}}+h^{\mathrm{B}}d_{\mathrm{B}}\|^2}{4N_0}\right)\right\}<\frac{16N_0^2}{\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}}\\
&-\frac{16\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}-\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}+\frac{4\pi N_0^2\Gamma(d_{\mathrm{A}},d_{\mathrm{B}})}{(\Upsilon_{\mathrm{A}}\Upsilon_{\mathrm{B}}+\Gamma^2(d_{\mathrm{A}},d_{\mathrm{B}}))^{\frac{3}{2}}}.
\end{aligned}\tag{63}$$

REFERENCES

[1] R. Ahlswede, N. Cai, S-Y. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[2] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical layer network coding," *Proc. 12th Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, Sep. 2006, pp. 358–365 .

[3] P. Popovski and H. Yomo, "Wireless network coding by amplify-and-forward for bi-directional traffic flows," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 16–18, Jan. 2007.

[4] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[5] S. Zhang and S. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 788–796, Jun. 2009.

[6] W. Nam, S-Y. Chung, and Y. Lee, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[7] S. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.

[8] M. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[9] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[10] L. Yang, T. Yang, J. Yuan, and J. An, "Achieving the near-capacity of two-way relay channels with modulation-coded physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5225–5239, Sep. 2015.

[11] T. Huang, T. Yang, J. Yuan, and I. Land, "Design of irregular repeat-accumulate coded physical-layer network coding for Gaussian two-way relay channels," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 897–909, Mar. 2013.

[12] A. Tanc, T. Duman, and C. Tepedelenlioglu, "Design of LDPC codes for two-way relay systems with physical-layer network coding," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 2356–2359, Dec. 2013.

[13] L. Chen, Y. Yakufu, X. Yuan, and Z. Sun, "Design of BICM-ID for two-way relay channels," in *Proc. 2015 Int. Workshop High Mobility Wirel. Commun.*, Xi'an, China, Oct. 2015, pp. 11–15.

[14] M. Reed, C. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1693–1699, Dec. 1998.

[15] M. Reed, *Iterative Receiver Techniques for Coded Multiple Access Communication Systems*, Ph.D. thesis, Universty of South Australia, Adelaide, SA, Australia, 1999.

[16] F. Brannstrom, T. Aulin, L. Rasmussen, and A. Grant, "Convergence analysis of iterative detectors for narrow-band multiple access," *Proc. IEEE Global Telecom. Conf.*, Taipei, Taiwan, R.O.C., Nov. 2002.

[17] Q. Yang and S. Liew, "Asynchronous convolutional-coded physical-layer network coidng," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 3, pp. 1380–1395, Mar. 2015.

[18] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, Jun. 2009.

[19] V. Namboodiri, K. Venugopal, and B. Rajan, "Physical layer network coding for two-way relaying with QAM," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5074–5086. Oct. 2013.

[20] T. Ferrett, M. Valenti, and D. Torrieri, "An iterative noncoherent relay receiver for the two-way relay channel," in *Proc. IEEE Int. Conf. Commun.*, Budapest, Hungary, Jun. 2013, pp. 5903–5908.

[21] Z. Chen and H. Liu, "Spectrum efficient coded modulation design for two-way relay channels," *IEEE J. Sel. Areas in Commun.*, vol, 32, no. 2, pp. 251–263, Feb. 2014.

[22] X. Li, S. Zhang, and G. Qian, "Mapping and coding design for channel coded physical-layer network coding," *Proc. 2013 Int. Workshop High Mob. Wirel. Commun.*, Shanghai, China, Nov. 2013.

[23] R. Chang, S. Lin, and W. Chung, "Symbol and bit mapping optimization for physical-layer network coding with pulse amplitude modulation," *IEEE Trans. Wirel. Commun.*, vol. 12, no. 8, pp. 3956–3967, Aug. 2013.

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

[25] A. Chindapol and J. Ritcey, "Design, analysis and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 944–957, May 2001.

[26] J. Tan and G. Stüber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.* vol. 42, no. 2, pp. 662–671, Mar. 2005.

[27] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.

[28] D. To and J. Choi, "Convolutional codes in two-way relay networks with physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2724–2729, Sep. 2010.

[29] C. Vitiello, S. Pfletschinger, and M. Luise, "Decoding options for trellis codes in the two-way relay channel," in*Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun.*, Darmstadt, Germany, Jun. 2013, pp. 380–384.

[30] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–945, May 1998.

[31] A. Fabregas, A. Martinez, and G. Caire, *Bit-Interleaved Coded Modulation, Foundations and Trends in Communications and Information Theory*, Breda, The Netherlands: Now Publisher, 2008.

[32] S. Pfletschinger, "A practical physical-layer network coding scheme for the uplink of the two-way relay channel," in *Proc. 45th Asilomar Conf.*, Pacific Grove, CA, USA, pp. 1997–2001, Nov. 2011.

[33] X. Yuan, T. Yang, and I. B. Collings, "Multiple-input multiple-output two-way relaying: A space-division approach," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6421–6440, Oct. 2013.

[34] L. Addison, *Statistical Signal Processing-detection, Estimation and Time Series Analysis*, Reading, MA, USA: Addison-Wesley, 1991.

Authors' biographies not available at the time of publication.