

Threshold-Based One-Bit Soft Forwarding for a Network Coded Multi-Source Single-Relay System

Jun Li, *Member, IEEE*, Zihuai Lin, *Senior Member, IEEE*, Branka Vucetic, *Fellow, IEEE*, Jin Xu, Ming Xiao, *Senior Member, IEEE*, and Wen Chen, *Senior Member, IEEE*

Abstract—In this paper, we propose a threshold-based one-bit soft forwarding (TOB-SF) protocol for a multi-source relaying system with network coding, where two sources communicate with the destination with the help of a relay. Specifically in the TOB-SF protocol, the relay calculates the log-likelihood ratio (LLR) value of each network coded symbol, compares this LLR value with a pre-optimized threshold, and determines whether to transmit or keep silent. We are interested in optimizing the TOB-SF protocol in fading channels, and consider both the uncoded and low-density parity check coded systems. In the uncoded system, we first derive the bit error rate (BER) expressions at the destination, based on which, we derive the optimal threshold. Then we theoretically prove that the system can achieve the full diversity gain by using this threshold. Further, we optimize the power allocation at the relay to achieve a higher coding gain. In the coded system, we first optimize the LLR threshold. Then we develop a methodology to track the BER evolution at the destination by using Gaussian approximations. Based on the BER evolution, we further optimize the power allocation at the relay which minimizes the system BER. Simulation results show that the proposed TOB-SF protocol outperforms other conventional relaying protocols in terms of error performance.

Index Terms—One-bit soft forwarding, log-likelihood ratio, wireless network coding, power allocation.

I. INTRODUCTION

RELAY technology is an efficient tool to combat channel fading and enhance transmission rate by exploiting spatial degree of freedom [1–3]. In wireless relay networks, one or more relays listen to the signals transmitted by the source, perform signal processing on the received signals, and forward the processed signals to the destination. By combining the signals from the source and the relays, the destination can thus take the advantage of spatial diversity and make a reliable decision on the source’s information.

Manuscript received August 6, 2013; revised December 9, 2013 and February 9, 2014. The editor coordinating the review of this paper and approving it for publication was V. Stankovic.

J. Li, Z. Lin, and B. Vucetic are with the School of Electrical and Information Engineering, The University of Sydney, NSW, 2006, Australia (e-mail: {jun.li, zihuai.lin, branka.vucetic}@sydney.edu.au).

J. Xu is with the Wireless Technology Pre-research Department, Product R&D System, ZTE, Shenzhen, China (e-mail: xu.jin7@zte.com.cn).

M. Xiao is with the School of Electrical Engineering, Royal Institute of Technology, Sweden (e-mail: ming.xiao@ee.kth.se).

W. Chen is with the Department of Electronic Engineering, Shanghai Jiao Tong University, China (e-mail: wenchen@sjtu.edu.cn).

This research work is supported by Australian Research Council Programs DP120100405 and LP110100110, by the Research Fund of ZTE Corporation, by National 973 Programs 2013CB329001 and 2012CB316106, by NSF China 61161130529 and 61328101, by STCSM Science and Technology Innovation Program 13510711200, and by SEU National Key Lab on Mobile Communications 2013D11.

Digital Object Identifier 10.1109/TCOMM.2014.030814.130621

Early publications mainly suggest two relaying protocols, namely, amplify-and-forward (AF) and decode-and-forward (DF) [2]. In the AF protocol, the relay amplifies the signals from the source, and forwards them to the destination. Although it is simple, the AF protocol suffers from noise propagations as the relay forwards noisy signals without processing. In the DF protocol, the relay detects/decodes the source’s messages and forwards them to the destination. By explicitly obtaining the source’s messages, the relay forwards recovered signals without noise. Another advantage of the DF protocol is that it can be jointly designed with channel codes [4–6] or wireless network coding [7, 8] to achieve a higher coding gain. However, if the relay makes a wrong decision, the error will propagate to the destination, causing a loss in diversity gain.

Channel-aware demodulators and link adaptive solutions are proposed to eliminate the error propagations in the conventional DF so as to achieve a full diversity gain [9–16]. Specifically, [9–13] propose novel network coding schemes to achieve the full diversity gain with the assumption that frame errors at the relay can be detected by using some CRC codes. If an error is detected in a frame, the relay will discard this frame when doing network coding. We have to mention that (1) by using CRC codes, extra check information will be transmitted, which increase the transmission delay; (2) by detecting and dropping the erroneous frames at the relay, some useful information inside the frame is discarded as well, which will reduce the coding gain of the system. In [14–16], channel-aware demodulators (i.e., some smart signal combining methods) at the destination are proposed to achieve the full diversity by weighing the signals from the sources and the relay. Intuitively, if the source-to-relay channel is worse than the relay-to-destination channel, the signal from the relay will multiply a smaller weight to alleviate the impact of possible error from the relay. However, instantaneous source-to-relay channel state information is needed.

Consequently, an advanced relay protocol by forwarding soft information, namely, estimate-and-forward (EF), is proposed to obtain a better error performance [17–20]. Rather than forwarding the received signals passively as the AF, or making hard decisions on the source’s messages aggressively as the DF, the relay in the EF protocol generates and forwards intermediate soft information. Usually, the soft information can be in the form of log-likelihood ratio (LLR) [17], soft bit [18], or soft mutual information [19, 20]. The EF protocol is shown to achieve a better error performance than the AF and

DF protocols. However, the EF protocol requires the relay to forward real values via wireless channels, which is impractical. Also, the optimal combining of the signals from the relay and the source is still an open problem. Sub-optimal combining causes performance loss, especially in the scenarios where there are multiple relays [21].

More practically, LLR-threshold based one-bit soft information forwarding (TOB-SF) protocol is proposed [22–24], where the relay calculates the LLR value of each received symbol transmitted by the source, and compares the LLR value with a pre-determined threshold. The relay forwards the hard decision of the corresponding symbol if the LLR value is larger than the threshold, and keeps silent otherwise. The TOB-SF protocol requires the same bandwidth as the DF protocol, while it achieves a better performance than the DF by reducing error propagations.

We note that the TOB-SF protocol is mainly investigated in one-way relay channels [22, 23]. In [24], a network coded TOB-SF protocol is proposed for two-way relay channels, which is a simple extension of [22] since a two-way relay channel, with perfect side information, can be decomposed into two one-way relay channels. However, these results are not directly applicable for general multi-source relay channels, where multiple sources communicate with their destination with the help of the relay.

Multi-source relay channels are common and basic building blocks in wireless networks, especially in cellular networks. Therefore, investigating efficient relaying protocols to control error propagations in such channels is an important issue [25–27]. In [25], the LLR values of network coded symbols are forwarded by the relay. As discussed above, transmitting LLR values over bandwidth-limited wireless channels is impractical. Also, the sub-optimal combining of the signals at the destination will lead to a performance loss. In [26], quantization schemes on the LLR values are proposed. However, the iterative quantization process is complex, and impractical to change a quantization codebook each time when the source-to-relay channels change. In addition, the quantization scheme of [26] is based on the maximization of the mutual information between the LLR and its quantized result, which cannot guarantee the optimal error performance at the destination. In [27], the authors extend the power scaling schemes for one-way relay channels [28] to multi-source relay channels. Although the full diversity gain is achieved, the error performance is not fully optimized, i.e., the coding gain can be further improved.

In this paper, we design a network coded TOB-SF protocol in a two-source relaying system over fading channels, and consider both the uncoded and low-density parity check (LDPC) [29, 30] coded systems. In the uncoded system, we first derive the bit error rate (BER) expressions at the destination and optimize the threshold by minimizing the BER. Then we theoretically prove that the system can achieve the full diversity gain by using the proposed threshold. Furthermore, we optimize the power allocation at the relay to achieve a higher coding gain. In the coded system with LDPC codes, we first optimize the threshold. Then we develop a methodology to track the BER evolutions in the iterative receiver at the destination, based on which, we further optimize the power allocation at the relay.

The theoretical contributions of this paper are as follows. In the uncoded system, we (a) derive the BER expressions at the destination, (b) derive the exact expression of the optimal threshold, which is in the form of $\ln \rho + c$, where ρ is the signal-to-noise ratio (SNR), and c is a constant to SNR, (c) prove theoretically that the system can achieve the full diversity gain by using the optimal threshold, and (d) optimize the power allocation at the relay to further enhance the error performance. In the coded system with LDPC codes, we (a) optimize the threshold, (b) develop a methodology to track the BER evolution of the iterative receiver at the destination by using Gaussian approximations, and (c) optimize the power allocation at the relay based on the BER evolution results. In the simulations, we use the DF, EF, and the power scaling scheme from [27] as benchmarks. Simulation results show that the proposed TOB-SF protocol outperforms the benchmarks in terms of the error performance.

II. SYSTEM MODEL

Consider an orthogonal uplink relay channel with two sources, one relay and one destination as shown in Fig. 1, where the two sources \mathcal{S}_1 and \mathcal{S}_2 broadcast their messages to the common destination \mathcal{D} with the help of a half-duplex relay \mathcal{R} . Each transmission period consists of three phases. In the first phase, \mathcal{S}_1 broadcasts its message, and in second phase, \mathcal{S}_2 broadcasts its message, to the relay and the destination. After the first two phases, the relay generates the network coded message based on the signals from the two sources, which is then forwarded to the destination during the third phase. At the end of each transmission period, the destination decodes the messages of the two sources based on the signals from the sources and the relay.

As three phases are needed in each transmission period, one may consider to take the advantage of the physical layer network coding (PNC) to reduce one transmission phase. In [31], a novel decoder has been proposed to solve the synchronization problem in the PNC. However, the assumption that the amplitudes of the two source-to-relay channels are identical makes the PNC inapplicable in our system, since we model the channels as random fading.

We denote $h_{i\mathcal{R}}$, $i = 1, 2$, $h_{i\mathcal{D}}$, and $h_{\mathcal{R}\mathcal{D}}$ as the channel coefficients between \mathcal{S}_i and \mathcal{R} , between \mathcal{S}_i and \mathcal{D} , and between \mathcal{R} and \mathcal{D} , respectively, as shown in Fig. 1, and denote $d_{i\mathcal{R}}$, $d_{i\mathcal{D}}$, and $d_{\mathcal{R}\mathcal{D}}$ as the distances between \mathcal{S}_i and \mathcal{R} , between \mathcal{S}_i and \mathcal{D} , and between \mathcal{R} and \mathcal{D} , respectively. We assume that $h_{i\mathcal{R}}$, $h_{i\mathcal{D}}$, and $h_{\mathcal{R}\mathcal{D}}$ are independent and identically Rayleigh distributed with the channel gains as $\lambda_{i\mathcal{R}}$, $\lambda_{i\mathcal{D}}$, and $\lambda_{\mathcal{R}\mathcal{D}}$, respectively. These channel gains are related to the corresponding distances with the attenuation exponent γ , i.e., $\lambda_{i\mathcal{R}} = 1/(d_{i\mathcal{R}})^\gamma$, $\lambda_{i\mathcal{D}} = 1/(d_{i\mathcal{D}})^\gamma$, and $\lambda_{\mathcal{R}\mathcal{D}} = 1/(d_{\mathcal{R}\mathcal{D}})^\gamma$. We consider quasi-static fading channels, i.e., the channel coefficients are constant during one transmission period, and change independently from one period to another.

Let us assume that each phase in a transmission period consists of l time slots, and thus each transmission period lasts $3l$ time slots. Each source \mathcal{S}_i transmits the binary phase-shift keying (BPSK) symbol vector $\mathbf{x}_i = (x_i^1, \dots, x_i^l)^T$, $x_i^j \in \{\pm 1\}$ and $j \in \{1, \dots, l\}$, with the power E_i . The

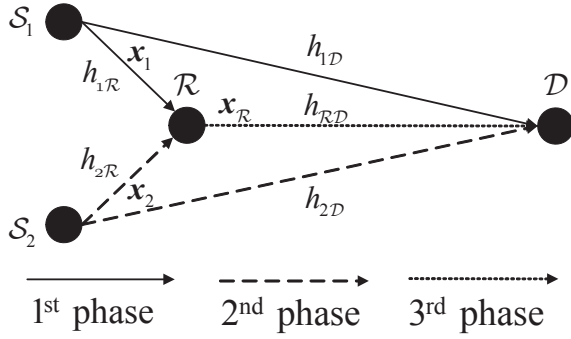


Fig. 1. The orthogonal uplink relay channel with two sources, one relay, and one destination.

received signals at the relay and at the destination from S_i can be expressed as

$$\begin{aligned} \mathbf{y}_{iR} &= h_{iR} \sqrt{E_i} \mathbf{x}_i + \mathbf{n}_{iR}, \\ \mathbf{y}_{iD} &= h_{iD} \sqrt{E_i} \mathbf{x}_i + \mathbf{n}_{iD}, \end{aligned} \quad (1)$$

respectively, where the vectors \mathbf{y}_{iR} and \mathbf{y}_{iD} consist of l received signals, i.e., $\mathbf{y}_{iR} = (y_{iR}^1, \dots, y_{iR}^l)^T$ and $\mathbf{y}_{iD} = (y_{iD}^1, \dots, y_{iD}^l)^T$, the vector $\mathbf{n}_{iR} = (n_{iR}^1, \dots, n_{iR}^l)^T$ consists of l additive white Gaussian noise (AWGN) samples at the relay, and the vector, $\mathbf{n}_{iD} = (n_{iD}^1, \dots, n_{iD}^l)^T$, consists of l AWGN samples at the destination. We assume that all the noise samples at the relay and the destination are with a mean zero and the same variance σ^2 , and the sources' power satisfies $E_1 = E_2 = 1$. We define the SNR as $\rho \triangleq 1/\sigma^2$.

After receiving \mathbf{y}_{iR} , the relay detects/decodes \mathbf{x}_i and generates the network coded message as $\mathbf{x}_R = (x_R^1, \dots, x_R^l)^T$, $x_R^j \in \{\pm 1\}$, based on the hard decisions of \mathbf{x}_i , denoted by $\hat{\mathbf{x}}_i$. Since the network coding operation, i.e., XOR, between two bits is equivalent to the multiplication of their BPSK symbols, a network coded symbol x_R^j can be obtained by $x_R^j = \hat{x}_1^j \hat{x}_2^j$, where \hat{x}_i^j is the hard decision of x_i^j at the relay. Hence, the network coded message \mathbf{x}_R is obtained by implementing inner product (vector-wise multiplication) on $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$. Note that as shown in [10], non-binary NC is necessary to achieve the full diversity gain in multiple-relay cases. However, since we only consider the one-relay case, we can use the conventional XOR based NC to achieve the full diversity, i.e., two-order diversity, in our system. Then we have $\mathbf{x}_R = \hat{\mathbf{x}}_1 \cdot \hat{\mathbf{x}}_2 = (\hat{x}_1^1 \hat{x}_2^1, \dots, \hat{x}_1^l \hat{x}_2^l)^T$. In our TOB-SF protocol, the relay only forwards those network coded symbols in \mathbf{x}_R whose absolute LLR values are larger than a preset threshold, and keeps silent otherwise. Then the received signal at the destination from the relay can be written as

$$\mathbf{y}_{RD} = h_{RD} \sqrt{\alpha v} \cdot \mathbf{x}_R + \mathbf{n}_{RD}, \quad (2)$$

where $0 < \alpha \leq 1$ is the power allocation at the relay, and $\mathbf{v} = (v_1, \dots, v_l)^T$ is an indicator vector, with $v_j = 1$ representing the transmission of x_R^j , and $v_j = 0$ representing being silent of the relay. Also in (2), $\mathbf{y}_{RD} = (y_{RD}^1, \dots, y_{RD}^l)^T$ is the received signal vector, and $\mathbf{n}_{RD} = (n_{RD}^1, \dots, n_{RD}^l)^T$ is the AWGN vector at the destination.

III. THRESHOLD-BASED ONE-BIT SOFT FORWARDING PROTOCOL

Recall that in the DF protocol, decision errors in \mathbf{x}_R are largely propagated to the destination from the relay, which leads to a severe performance degradation. On the other hand, the EF protocol mitigates the error propagations by forwarding the soft information of each symbol. However, it is quite impractical of the EF protocol to transmit a real number instead of a single bit along the R -to- D channel. The TOB-SF protocol combines the advantages of both protocols by alleviating error propagations with simple one-bit transmissions.

In the TOB-SF protocol, we consider both uncoded system and coded system. For the uncoded system, there is no channel coding at each source, while for the coded system, LDPC codes are applied to the two sources. Without a loss of generality, we focus on the j -th symbol in \mathbf{x}_i , i.e., x_i^j . Based on the detection/decoding results of x_1^j and x_2^j , the relay calculates the LLR value for each network coded symbol x_R^j in \mathbf{x}_R . We denote $L_{x_1^j, R}$ as the LLR value of x_1^j after detection/decoding at the relay, and denote $L_{x_2^j, R}$ as the LLR value of x_2^j . Then we have

$$L_{x_R^j, R} = 2 \tanh^{-1} \left(\tanh \left(\frac{L_{x_1^j, R}}{2} \right) \tanh \left(\frac{L_{x_2^j, R}}{2} \right) \right), \quad (3)$$

where $\tanh(z) = \frac{\exp(2z)-1}{\exp(2z)+1}$ is the hyperbolic function. In the uncoded system, $L_{x_i^j, R}$ is calculated based on the conditional probability density function (PDF) of y_{iR}^j , i.e.,

$$\begin{aligned} L_{x_i^j, R} &= \ln \frac{p(y_{iR}^j | x_i^j = 1, h_{iR})}{p(y_{iR}^j | x_i^j = -1, h_{iR})} \\ &= \frac{2h_{iR}}{\sigma^2} y_{iR}^j. \end{aligned} \quad (4)$$

For the LDPC coded system, $L_{x_i^j, R}$ can be obtained from the output of the decoder at the relay.

After obtaining $L_{x_R^j, R}$ based on (4), the relay compares $|L_{x_R^j, R}|$ with a preset threshold L_T ($L_T \geq 0$). If $|L_{x_R^j, R}|$ is larger than L_T , the j -th indicator v_j in the vector \mathbf{v} is set to one, otherwise it is set to zero, i.e.,

$$v_j = \begin{cases} 1, & |L_{x_R^j, R}| > L_T, \\ 0, & |L_{x_R^j, R}| \leq L_T. \end{cases} \quad (5)$$

By applying the indicator vector \mathbf{v} , the relay either transmits BPSK symbols or keeps silent.

Practically, the destination does not know the indicator vector \mathbf{v} at the relay. However, the destination can detect whether a symbol is transmitted from the relay by monitoring the power of the received signal. When the received SNR from the relay is large, the destination can always make a correct detection. To simplify the following analysis, we assume that the destination can perfectly detect the relay's status. When the destination makes a decision that relay is silent, it will only adopts the signals from the sources.

In the following, we will discuss the optimizations of the threshold L_T and the power allocation α at the relay to enhance the BER performance at the destination. Practically, we assume that the instantaneous channel state information

(CSI) of $h_{1\mathcal{R}}$ and $h_{2\mathcal{R}}$, and the statistical CSI of $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$ are available at the relay during the optimizations. The extension from BPSK to a higher order modulation is possible, where multiple thresholds need to be set and jointly optimized to minimize the BER. However, since binary NC and binary LDPC code are considered in the following analysis, we only adopt the BPSK modulation in our paper.

IV. OPTIMIZATIONS AND PERFORMANCE ANALYSIS FOR UNCODED SYSTEM

A. Threshold Optimization

We denote \mathcal{E}_c as the event that $|L_{x_{\mathcal{R}}^j, \mathcal{R}}|$ is larger than L_T and $x_{\mathcal{R}}^j$ is correct, denote \mathcal{E}_e as the event that $|L_{x_{\mathcal{R}}^j, \mathcal{R}}|$ is larger than L_T and $x_{\mathcal{R}}^j$ is in error, and denote \mathcal{E}_s as the event that $|L_{x_{\mathcal{R}}^j, \mathcal{R}}|$ is no larger than L_T , i.e.,

$$\begin{aligned} \mathcal{E}_c &: |L_{x_{\mathcal{R}}^j, \mathcal{R}}| > L_T, \text{ and } x_{\mathcal{R}}^j = x_1^j x_2^j, \\ \mathcal{E}_e &: |L_{x_{\mathcal{R}}^j, \mathcal{R}}| > L_T, \text{ and } x_{\mathcal{R}}^j \neq x_1^j x_2^j, \\ \mathcal{E}_s &: |L_{x_{\mathcal{R}}^j, \mathcal{R}}| \leq L_T. \end{aligned} \quad (6)$$

Given a set of channel realizations $h_{1\mathcal{R}}$, $h_{2\mathcal{R}}$, $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$, we denote P_e as the average BER of the two sources at the destination. We have $P_e = \frac{1}{2}(P_{e,1} + P_{e,2})$, where $P_{e,i}$ is the BER of \mathcal{S}_i at the destination. We also denote $P_{i,mrc}$, $P_{i,neg}$, and $P_{i,dir}$ as the error probabilities of a symbol x_i^j at the destination given that the relay forwards the correct $x_{\mathcal{R}}^j$ to the destination, forwards the incorrect $x_{\mathcal{R}}^j$ to the destination, and keeps silent, respectively. Note that $P_{i,mrc}$, $P_{i,neg}$, and $P_{i,dir}$ rely on the channels $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$, while $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$ rely on the channels $h_{1\mathcal{R}}$ and $h_{2\mathcal{R}}$. When averaged over $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$, we have

$$\begin{aligned} \mathbb{E}_{h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{RD}}}(P_{e,i}) &= \mathbb{E}_{h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{RD}}}(P_{i,mrc}) \Pr(\mathcal{E}_c) + \\ & \mathbb{E}_{h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{RD}}}(P_{i,neg}) \Pr(\mathcal{E}_e) + \mathbb{E}_{h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{RD}}}(P_{i,dir}) \Pr(\mathcal{E}_s). \end{aligned} \quad (7)$$

For the sake of simplicity, we use $\mathbb{E}(\cdot)$ instead of $\mathbb{E}_{h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{RD}}}(\cdot)$ in the following analysis.

We focus on the calculations of $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$. First, we investigate the PDF of the LLR value $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ at the relay. For a given channel realization, $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ can be approximated as a Gaussian variable, whose variance is twice the absolute value of its mean [32]. Therefore, the PDF of $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ can be written as

$$p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) = \frac{1}{\sqrt{2\pi\sigma_{L\oplus}^2}} \exp\left(-\frac{(L - m_{L\oplus})^2}{2\sigma_{L\oplus}^2}\right), \quad (8)$$

where $m_{L\oplus}$ is the mean value, $\sigma_{L\oplus}^2$ is the variance of the $L_{x_{\mathcal{R}}^j, \mathcal{R}}$, and we have $\sigma_{L\oplus}^2 = 2|m_{L\oplus}|$.

Then we determine $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$, which depend on the PDF $p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L)$ and the threshold L_T . Without

a loss of generality, we consider $x_1^j x_2^j = 1$, and we have

$$\begin{aligned} \Pr(\mathcal{E}_c) &= \int_{L_T}^{+\infty} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL, \\ \Pr(\mathcal{E}_e) &= \int_{-\infty}^{-L_T} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL, \\ \Pr(\mathcal{E}_s) &= \int_{-L_T}^{L_T} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL \\ &= 1 - \Pr(\mathcal{E}_c) - \Pr(\mathcal{E}_e). \end{aligned} \quad (9)$$

At the destination, the average BER P_e can be minimized by optimizing the threshold L_T . The following theorem offers the solution of this optimization problem.

Theorem 1: When the statistical CSI of $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$ is available at the relay, the optimal threshold that minimizes the BER at the destination can be written as

$$L_T^* = \ln \frac{\mathbb{E}(P_{neg}) - \mathbb{E}(P_{dir})}{\mathbb{E}(P_{dir}) - \mathbb{E}(P_{mrc})}, \quad (10)$$

where $\mathbb{E}(P_{mrc}) = \frac{1}{2}(\mathbb{E}(P_{1,mrc}) + \mathbb{E}(P_{2,mrc}))$, $\mathbb{E}(P_{neg}) = \frac{1}{2}(\mathbb{E}(P_{1,neg}) + \mathbb{E}(P_{2,neg}))$, and $\mathbb{E}(P_{dir}) = \frac{1}{2}(\mathbb{E}(P_{1,dir}) + \mathbb{E}(P_{2,dir}))$.

Proof: Please refer to Appendix A. ■

B. Analysis on $\mathbb{E}(P_{i,mrc})$, $\mathbb{E}(P_{i,neg})$, and $\mathbb{E}(P_{i,dir})$

We can see from (10) that the optimal threshold L_T^* is based on $P_{i,mrc}$, $P_{i,neg}$, and $P_{i,dir}$ at the destination, which are related to the channels $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$. Now, we will determine these three probabilities when the statistical CSI of $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{RD}}$ is available at the relay.

We assume that LLR combining is applied at the destination to detect the symbol x_i^j . We denote $L_{x_i^j, \mathcal{D}}$ and $L_{x_{\mathcal{R}}^j, \mathcal{D}}$ as the received LLR values at the destination for x_i^j and $x_{\mathcal{R}}^j$, respectively, and denote $L_{x_i^j, \mathcal{D}, ext}$ as the extrinsic LLR for x_i^j from the network coding between x_i^j and $x_{\mathcal{R}}^j$, where $\bar{i} = 1, 2$ and $\bar{i} \neq i$. The combined LLR for x_i^j , denoted by $L_{x_i^j, \mathcal{D}, comb}$ is calculated as

$$\begin{aligned} L_{x_i^j, \mathcal{D}, comb} &= L_{x_i^j, \mathcal{D}} + L_{x_i^j, \mathcal{D}, ext} \\ &= L_{x_i^j, \mathcal{D}} + 2 \tanh^{-1} \left(\tanh \left(\frac{L_{x_i^j, \mathcal{D}}}{2} \right) \tanh \left(\frac{L_{x_{\mathcal{R}}^j, \mathcal{D}}}{2} \right) \right). \end{aligned} \quad (11)$$

In (11), given $h_{i\mathcal{D}}$ and $h_{\mathcal{RD}}$, the received LLRs at the destination, i.e., $L_{x_i^j, \mathcal{D}}$ and $L_{x_{\mathcal{R}}^j, \mathcal{D}}$, are Gaussian distributed with their variances being twice the absolute values of their means [33]. In addition, the LLR $L_{x_i^j, \mathcal{D}, ext}$ from the network coding process can be approximated as a Gaussian variable with its variance being twice the absolute value of its mean [32].

Denote by $m_{L_{i\mathcal{D}}} \triangleq \frac{2h_{i\mathcal{D}}^2 x_i^j}{\sigma^2}$ the mean value of $L_{x_i^j, \mathcal{D}}$, and by $m_{L_{\mathcal{RD}}} \triangleq \frac{2h_{\mathcal{RD}}^2 \alpha v_j^2 x_{\mathcal{R}}^j}{\sigma^2}$ the mean value of $L_{x_{\mathcal{R}}^j, \mathcal{D}}$. Then the mean value of $L_{x_i^j, \mathcal{D}, ext}$, denoted by $m_{L_{i\mathcal{D}, ext}}$, can be calculated by utilizing the function $\phi(\cdot)$ defined in [32], i.e.,

$$\begin{aligned} m_{L_{i\mathcal{D}, ext}} &= x_i^j x_{\mathcal{R}}^j \phi^{-1}(\phi(|m_{L_{i\mathcal{D}}}|) + \\ & \phi(|m_{L_{\mathcal{RD}}}|) - \phi(|m_{L_{i\mathcal{D}}}|)\phi(|m_{L_{\mathcal{RD}}}|)), \end{aligned} \quad (12)$$

where the function $\phi(\cdot)$ is expressed as

$$\phi(z) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi z}} \int_{-\infty}^{\infty} \tanh\left(\frac{u}{2}\right) \exp\left(-\frac{(u-z)^2}{4z}\right) du, & z > 0, \\ 1, & z = 0. \end{cases} \quad (13)$$

When $z > 0$, the function $\phi(\cdot)$ is bounded as [32]

$$\sqrt{\frac{\pi}{z}} \exp\left(-\frac{z}{4}\right) \left(1 - \frac{3}{z}\right) < \phi(z) < \sqrt{\frac{\pi}{z}} \exp\left(-\frac{z}{4}\right) \left(1 + \frac{1}{7z}\right). \quad (14)$$

We can see from (14) that when z is large enough, the two bounds converge to $\sqrt{\frac{\pi}{z}} \exp\left(-\frac{z}{4}\right)$. Therefore, given z_1 and z_2 , when both of them are large enough, we have

$$\phi(z_1) + \phi(z_2) - \phi(z_1)\phi(z_2) \approx \max(\phi(z_1), \phi(z_2)). \quad (15)$$

Similarly, we can approximate $m_{L_{iD},ext}$ in (12) in the high SNR region as

$$\begin{aligned} m_{L_{iD},ext} &\approx x_i^j x_{\mathcal{R}}^j \phi^{-1}(\max(\phi(|m_{L_{iD}}|), \phi(|m_{L_{\mathcal{R}D}}|))) \\ &= \frac{2x_i^j x_{\mathcal{R}}^j}{\sigma^2} \min(h_{iD}^2, h_{\mathcal{R}D}^2 \alpha v_j^2). \end{aligned} \quad (16)$$

As we have $L_{x_i^j, \mathcal{D}, comb} = L_{x_i^j, \mathcal{D}} + L_{x_i^j, \mathcal{D}, ext}$, the combined LLR $L_{x_i^j, \mathcal{D}, comb}$ can be viewed as a Gaussian variable with the mean $m_{L_{iD}} + m_{L_{iD}, ext}$ and variance $2(|m_{L_{iD}}| + |m_{L_{iD}, ext}|)$. Given the channel coefficients h_{1D} , h_{2D} , and $h_{\mathcal{R}D}$, the instantaneous error probabilities can be obtained by utilizing $Q(\cdot)$ function, where $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} \exp\left(-\frac{u^2}{2}\right) du$. Specifically, when \mathcal{E}_c happens to the symbol $x_{\mathcal{R}}^j$, we have $x_i^j = x_{\mathcal{R}}^j$. In this case, $m_{L_{iD}}$ and $m_{L_{iD}, ext}$ have the same sign, and thus $|m_{L_{iD}} + m_{L_{iD}, ext}| = |m_{L_{iD}}| + |m_{L_{iD}, ext}|$. Therefore, we obtain the error probability of x_i^j as

$$P_{i, mrc} = Q\left(\sqrt{\frac{|m_{L_{iD}}| + |m_{L_{iD}, ext}|}{2}}\right). \quad (17)$$

When \mathcal{E}_e happens, we have $x_i^j = -x_{\mathcal{R}}^j$. In this case, $m_{L_{iD}}$ and $m_{L_{iD}, ext}$ have the opposite signs, and we have $|m_{L_{iD}} + m_{L_{iD}, ext}| = ||m_{L_{iD}}| - |m_{L_{iD}, ext}||$. The calculation of $P_{i, neg}$ depends on the relation between $|m_{L_{iD}}|$ and $|m_{L_{iD}, ext}|$, i.e.,

$$P_{i, neg} = \begin{cases} Q\left(\sqrt{\frac{(|m_{L_{iD}}| - |m_{L_{iD}, ext}|)^2}{2(|m_{L_{iD}}| + |m_{L_{iD}, ext}|)}}\right), & |m_{L_{iD}}| > |m_{L_{iD}, ext}|, \\ 1 - Q\left(\sqrt{\frac{(|m_{L_{iD}}| - |m_{L_{iD}, ext}|)^2}{2(|m_{L_{iD}}| + |m_{L_{iD}, ext}|)}}\right), & |m_{L_{iD}}| \leq |m_{L_{iD}, ext}|. \end{cases} \quad (18)$$

When \mathcal{E}_s happens, we have $v_j = 0$. Thus, the mean value $m_{L_{iD}, ext}$ equals zero, and we obtain

$$P_{i, dir} = Q\left(\sqrt{\frac{|m_{L_{iD}}|}{2}}\right). \quad (19)$$

The statistical CSI based error probabilities can be obtained by averaging $P_{i, mrc}$, $P_{i, neg}$, and $P_{i, dir}$ over h_{1D}^2 , h_{2D}^2 , and $h_{\mathcal{R}D}^2$, where the PDFs of h_{iD}^2 and $h_{\mathcal{R}D}^2$ are $p_{h_{iD}^2}(h) = \frac{1}{\lambda_{iD}} \exp\left(-\frac{h}{\lambda_{iD}}\right)$ and $p_{h_{\mathcal{R}D}^2}(h) = \frac{1}{\lambda_{\mathcal{R}D}} \exp\left(-\frac{h}{\lambda_{\mathcal{R}D}}\right)$, respectively. Alternatively, by defining $m_{L_{iD}, mrc} \triangleq |m_{L_{iD}}| + |m_{L_{iD}, ext}|$, by defining $m_{L_{iD}, neg} \triangleq \frac{(|m_{L_{iD}}| - |m_{L_{iD}, ext}|)^2}{|m_{L_{iD}}| + |m_{L_{iD}, ext}|}$,

and by defining $m_{L_{iD}, dir} \triangleq |m_{L_{iD}}|$, we average $P_{i, mrc}$, $P_{i, neg}$, and $P_{i, dir}$ over the variables $m_{L_{iD}, mrc}$, $m_{L_{iD}, neg}$, and $m_{L_{iD}, dir}$, respectively. The PDFs of these three variables are given in the following lemma.

Lemma 1: The PDF of $m_{L_{iD}, mrc}$ can be expressed as

$$p_{m_{L_{iD}, mrc}}(m) = \frac{\lambda_i \bar{\lambda}_i}{\lambda_i - \bar{\lambda}_i} (\exp(-\bar{\lambda}_i m) - \exp(-\lambda_i m)), \quad (20)$$

where $\lambda_i = \frac{1}{\sigma^2 \lambda_{iD}}$ and $\bar{\lambda}_i = \frac{1}{\sigma^2 \lambda_{iD}} + \frac{1}{\sigma^2 \lambda_{\mathcal{R}D}}$. The PDF of $m_{L_{iD}, neg}$ can be expressed as

$$\begin{aligned} p_{m_{L_{iD}, neg}}(m) &= \frac{\lambda_i^2 \bar{\lambda}_i^2}{\lambda_i^2 - \bar{\lambda}_i^2} \int_0^{\infty} \frac{u^2}{\sqrt{m}} (\exp(-\bar{\lambda}_i u^2 - \lambda_i \sqrt{m} u) \\ &\quad - \exp(-\lambda_i u^2 - \bar{\lambda}_i \sqrt{m} u)) du. \end{aligned} \quad (21)$$

The PDF of $m_{L_{iD}, dir}$ can be expressed as

$$p_{m_{L_{iD}, dir}}(m) = \lambda_i \exp(-\lambda_i m). \quad (22)$$

Proof: Please refer to Appendix B. ■

Based on Lemma 1, we can derive the statistical CSI based error probabilities, i.e., the expectations of $P_{i, mrc}$, $P_{i, neg}$, and $P_{i, dir}$. Specifically, the expectation of $P_{i, mrc}$ is calculated as

$$\begin{aligned} \mathbb{E}(P_{i, mrc}) &= \frac{\lambda_i}{\lambda_i - \bar{\lambda}_i} \int_0^{\infty} Q\left(\sqrt{\frac{m}{2}}\right) \bar{\lambda}_i \exp(-\bar{\lambda}_i m) dm - \\ &\quad \frac{\bar{\lambda}_i}{\lambda_i - \bar{\lambda}_i} \int_0^{\infty} Q\left(\sqrt{\frac{m}{2}}\right) \lambda_i \exp(-\lambda_i m) dm \\ &= \frac{\lambda_i}{2(\lambda_i - \bar{\lambda}_i)} \left(1 - \sqrt{\frac{1}{1 + 4\bar{\lambda}_i}}\right) - \\ &\quad \frac{\bar{\lambda}_i}{2(\lambda_i - \bar{\lambda}_i)} \left(1 - \sqrt{\frac{1}{1 + 4\lambda_i}}\right). \end{aligned} \quad (23)$$

The expectation of $P_{i, neg}$ is calculated as follows. First, we calculate the probability that $|m_{L_{iD}}| > |m_{L_{iD}, ext}|$, i.e., $\Pr(|m_{L_{iD}}| > |m_{L_{iD}, ext}|)$. From Appendix B, we have

$$\Pr(|m_{L_{iD}}| > |m_{L_{iD}, ext}|) = \frac{\bar{\lambda}_i}{\lambda_i + \bar{\lambda}_i}. \quad (24)$$

Then we obtain the expectation of $P_{i, neg}$ as [22]

$$\begin{aligned} \mathbb{E}(P_{i, neg}) &= \int_0^{\infty} Q\left(\sqrt{\frac{m}{2}}\right) p_{m_{L_{iD}, neg}}(m) dm \Pr(|m_{L_{iD}}| > |m_{L_{iD}, ext}|) \\ &\quad + \left(1 - \int_0^{\infty} Q\left(\sqrt{\frac{m}{2}}\right) p_{m_{L_{iD}, neg}}(m) dm\right) \\ &\quad \Pr(|m_{L_{iD}}| \leq |m_{L_{iD}, ext}|) \\ &= \frac{\lambda_i}{\lambda_i + \bar{\lambda}_i} - \frac{\lambda_i - \bar{\lambda}_i}{\lambda_i + \bar{\lambda}_i} \int_0^{\infty} Q\left(\sqrt{\frac{m}{2}}\right) p_{m_{L_{iD}, neg}}(m) dm \\ &= \frac{\lambda_i}{\lambda_i + \bar{\lambda}_i} - \frac{\lambda_i^2 \bar{\lambda}_i^2}{(\lambda_i + \bar{\lambda}_i)^2} \left(f\left(\frac{1}{4}, \bar{\lambda}_i, \lambda_i\right) - f\left(\frac{1}{4}, \lambda_i, \lambda_i\right)\right), \end{aligned} \quad (25)$$

where the function $f(\cdot)$ is defined in Equation (A.4) in [22].

The expectation of $P_{i,dir}$ is calculated as

$$\begin{aligned}\mathbb{E}(P_{i,dir}) &= \int_0^\infty Q\left(\sqrt{\frac{m}{2}}\right) \lambda_i \exp(-\lambda_i m) dm \\ &= \frac{1}{2} \left(1 - \sqrt{\frac{1}{1+4\lambda_i}}\right).\end{aligned}\quad (26)$$

Hence, when the statistical CSI of h_{1D} , h_{2D} , and h_{RD} is available at the relay, we can obtain the optimal threshold L_T^* based on (10) with $\mathbb{E}(P_{i,mrc})$, $\mathbb{E}(P_{i,neg})$, and $\mathbb{E}(P_{i,dir})$.

C. Diversity Gain

We now investigate the diversity order of the system with our TOB-SF protocol. When the SNR $\rho = \frac{1}{\sigma^2}$ is large enough, we make approximations on $\mathbb{E}(P_{i,mrc})$, $\mathbb{E}(P_{i,neg})$, and $\mathbb{E}(P_{i,dir})$. First, since the function $f(\cdot)$ in (25) is proportional to ρ^{-1} [22], it is straightforward to obtain the approximation of $\mathbb{E}(P_{i,neg})$ from (25), i.e.,

$$\begin{aligned}\mathbb{E}(P_{i,neg}) &\approx \frac{\lambda_i}{\lambda_i + \bar{\lambda}_i} \\ &= \frac{\lambda_{iD} \lambda_{RD}}{\lambda_{iD} \lambda_{RD} + \lambda_{iD} \lambda_{RD} + \frac{1}{\alpha} \lambda_{iD} \lambda_{iD}}.\end{aligned}\quad (27)$$

Then by utilizing the property of the function $Q(\cdot)$ shown in [34], we have

$$\begin{aligned}\mathbb{E}(P_{i,mrc}) &\approx \frac{3}{4\lambda_{iD} \frac{\alpha \lambda_{iD} \lambda_{RD}}{\lambda_{iD} + \alpha \lambda_{RD}}} \rho^{-2}, \\ \mathbb{E}(P_{i,dir}) &\approx \frac{1}{2\lambda_{iD}} \rho^{-1}.\end{aligned}\quad (28)$$

Based on (27) and (28), when ρ is large enough, we can approximate the statistical CSI based optimal threshold as

$$\begin{aligned}L_T^* &\approx \ln \frac{\mathbb{E}(P_{1,neg}) + \mathbb{E}(P_{2,neg})}{\mathbb{E}(P_{1,dir}) + \mathbb{E}(P_{2,dir})} \\ &\approx \ln \rho + \ln \frac{2\lambda_{1D} \lambda_{2D} \lambda_{RD}}{\lambda_{1D} \lambda_{RD} + \lambda_{2D} \lambda_{RD} + \frac{1}{\alpha} \lambda_{1D} \lambda_{2D}}.\end{aligned}\quad (29)$$

By defining $c \triangleq \ln \frac{2\lambda_{1D} \lambda_{2D} \lambda_{RD}}{\lambda_{1D} \lambda_{RD} + \lambda_{2D} \lambda_{RD} + \frac{1}{\alpha} \lambda_{1D} \lambda_{2D}}$, we can see that c does not depend on ρ , but depends on the channel gains and the power allocation α . Hence, we have $\lim_{\rho \rightarrow \infty} L_T^* = \ln \rho$.

We now focus on the diversity gain of each source with the optimal threshold L_T^* . Since we have $\lim_{\rho \rightarrow \infty} L_T^* = \ln \rho$, we only need to investigate the diversity gain with the threshold $\ln \rho$.

Theorem 2: By utilizing the TOB-SF protocol with the threshold $\ln \rho$, each source can achieve the full diversity gain of the system, i.e., a diversity of two.

Proof: Please refer to Appendix C. ■

From the proof of *Theorem 2*, we can see that by using the threshold $\ln \rho$, the probability $\mathbb{E}(\Pr(\mathcal{E}_e))$ is proportional to ρ^{-2} , and $\mathbb{E}(\Pr(\mathcal{E}_s))$ is proportional to ρ^{-1} . Considering the probabilities in (27) and (28), we can obtain that the BER for \mathcal{S}_i , i.e., $P_{e,i}$, is proportional to ρ^{-2} when averaged over all channel parameters h_{1R} , h_{2R} , h_{1D} , h_{2D} , and h_{RD} . Thus, each source can achieve an order of two diversity gain by using the threshold $\ln \rho$.

Remark 1: The optimal threshold L_T^* can be approximated in the form of $\ln \rho + c$, where $\ln \rho$ is the key to guaranteeing the full diversity gain, as shown in the proof of *Theorem 2*, and c is the key to achieving the optimal coding gain.

D. Power Allocation

In this subsection, we will further enhance the error performance by optimizing the power allocation α at the relay. Specifically, by averaging $P_{e,i}$ over the channels h_{1D} , h_{2D} , and h_{RD} , we have $\mathbb{E}(P_{e,i}) = \mathbb{E}(P_{i,mrc}) \Pr(\mathcal{E}_c) + \mathbb{E}(P_{i,neg}) \Pr(\mathcal{E}_e) + \mathbb{E}(P_{i,dir}) \Pr(\mathcal{E}_s)$. We minimize the average BER, i.e., $\mathbb{E}(P_e) = \frac{1}{2}(\mathbb{E}(P_{e,1}) + \mathbb{E}(P_{e,2}))$, by optimizing α , and obtain the optimal α by solving $\frac{\partial \mathbb{E}(P_e)}{\partial \alpha} = 0$ with the constraint $0 < \alpha \leq 1$.

However, from the expression of the optimal threshold L_T^* in (29), we can see that L_T^* is a function of α . If L_T^* is used as the threshold, it is difficult to obtain the closed form of the optimal α when solving $\frac{\partial \mathbb{E}(P_e)}{\partial \alpha} = 0$. Fortunately, from the previous subsection, we know that $\ln \rho$ can serve as a sub-optimal threshold to achieve the full diversity gain. Since $\ln \rho$ is independent of α , the items $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$ are constants relative to α if we use $\ln \rho$ as the threshold. Hence, the derivation of the optimal α can be simplified and a closed form of the optimal α is obtained. In the sequel, we optimize α by assuming that $\ln \rho$ is applied as the threshold.

With the power constraint $0 < \alpha \leq 1$, by utilizing the approximations in (27) and (28), we derive the optimal power allocation, denoted by α^* , at the relay as follows. First, we relax the constraint of α , and let $\frac{\partial \mathbb{E}(P_e)}{\partial \alpha} = 0$, which lead to two solutions for α , i.e.,

$$\begin{aligned}\alpha_1 &= \frac{\lambda_{1D} \lambda_{2D}}{\sqrt{\frac{4\Pr(\mathcal{E}_e)}{3\Pr(\mathcal{E}_c)} \rho \lambda_{1D} \lambda_{2D} \lambda_{RD}} - \lambda_{1D} \lambda_{RD} - \lambda_{2D} \lambda_{RD}}, \\ \alpha_2 &= \frac{\lambda_{1D} \lambda_{2D}}{-\sqrt{\frac{4\Pr(\mathcal{E}_e)}{3\Pr(\mathcal{E}_c)} \rho \lambda_{1D} \lambda_{2D} \lambda_{RD}} - \lambda_{1D} \lambda_{RD} - \lambda_{2D} \lambda_{RD}}.\end{aligned}\quad (30)$$

Note that from (30), we have $\alpha_2 < 0$ and $\alpha_2 < \alpha_1$. Also, α_1 can be a negative value or a positive value, depending on $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, ρ , and the channel gains.

Then we investigate the optimal power allocation α^* based on the two value ranges of α_1 , i.e., $\alpha_1 \leq 0$ and $\alpha_1 > 0$. When $\alpha_1 \leq 0$, we first relax the constraint of α and have the following two facts. (i) $\mathbb{E}(P_e)$ is a monotonically increasing function (MIF) of α if $\alpha_2 < \alpha \leq \alpha_1$, and (ii) $\mathbb{E}(P_e)$ is a monotonically decreasing function (MDF) of α if $\alpha \leq \alpha_2$ or $\alpha > \alpha_1$. Then we consider the optimal α with the constraint $0 < \alpha \leq 1$. We can see that within the range $0 < \alpha \leq 1$, $\mathbb{E}(P_e)$ is an MDF of α . Therefore, the optimal power allocation can be obtained as $\alpha^* = 1$ when $\alpha_1 \leq 0$.

When $\alpha_1 > 0$, we first relax the constraint of α and have the following two facts. (i) $\mathbb{E}(P_e)$ is an MDF of α if $\alpha_2 < \alpha \leq \alpha_1$, and (ii) $\mathbb{E}(P_e)$ is an MIF of α if $\alpha \leq \alpha_2$ or $\alpha > \alpha_1$. Then we optimize α with the constraint $0 < \alpha \leq 1$. We can see that if $\alpha_1 > 1$, $\mathbb{E}(P_e)$ is an MDF of α in the range $0 < \alpha \leq 1$. Thus the optimal power allocation can be obtained as $\alpha^* = 1$ when $\alpha_1 > 1$. Also, if $0 < \alpha_1 \leq 1$, $\mathbb{E}(P_e)$ is an MDF of α when $0 < \alpha \leq \alpha_1$, and is an MIF of α when $\alpha_1 < \alpha \leq 1$. Thus, the optimal power allocation can be obtained as $\alpha^* = \alpha_1$ when $0 < \alpha_1 \leq 1$. With above discussions, the optimal power allocation α^* at the relay is given as

$$\alpha^* = \begin{cases} 1, & \alpha_1 \leq 0 \text{ or } \alpha_1 > 1, \\ \alpha_1, & 0 < \alpha_1 \leq 1. \end{cases}\quad (31)$$

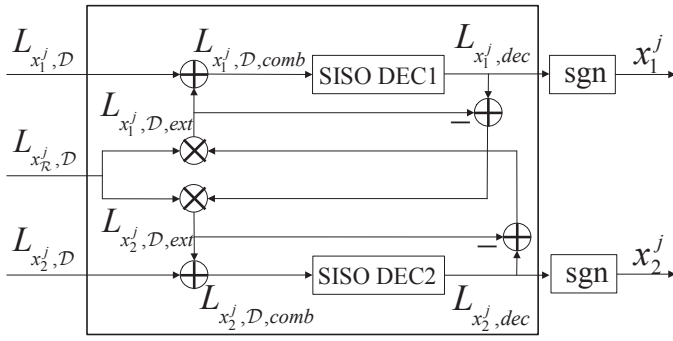


Fig. 2. The SISO iterative receiver at the destination. Note that ‘ \oplus ’ represents the LLR combining by addition, and ‘ \otimes ’ represents the LLR combining by network coding. SISO DEC $_i$ represents the SISO decoder for the source S_i .

Till now, we have optimized the threshold and power allocation for the uncoded system. In the next section, we will extend these optimizations to the LDPC coded system.

V. EXTENSIONS TO LDPC CODED SYSTEM

For the coded system, we assume that LDPC codes are applied at the two sources. Therefore, we can view x_i as an LDPC codeword of length l . The relay first decodes the two codewords of the two sources, and obtains the decoded codewords \hat{x}_1 and \hat{x}_2 . Then network coding is implemented on the two decoded codewords to obtain $x_{\mathcal{R}}$.

At the destination, a soft-input soft-output (SISO) iterative receiver is designed as shown in Fig. 2. We assume that each SISO decoder performs the sum-product decoding algorithm. As defined in the previous section, $L_{x_i^j, \mathcal{D}}$ and $L_{x_{\mathcal{R}}^j, \mathcal{D}}$ in Fig. 2 are the received LLR values at the destination from the source S_i and the relay, respectively. The extrinsic LLR $L_{x_i^j, \mathcal{D}, ext}$ for x_i^j is from the network coding between the symbols x_i^j and $x_{\mathcal{R}}^j$. The combined LLR $L_{x_i^j, \mathcal{D}, comb}$ is utilized as the input of S_i 's decoder. The output LLR of S_i 's decoder is denoted by $L_{x_i^j, dec}$. At the end of the final iteration, hard decision of x_i^j is made based on the sign of $L_{x_i^j, dec}$.

From Fig. 2 we can see that $L_{x_i^j, \mathcal{D}, ext}$ is updated during each iteration. Specifically, in the first iteration, there is no output from the SISO decoders, and $L_{x_i^j, \mathcal{D}, ext}$ is obtained by following the same method as that in the uncoded system, i.e., $L_{x_i^j, \mathcal{D}, ext}$ is calculated based on the received LLRs $L_{x_i^j, \mathcal{D}}$ and $L_{x_{\mathcal{R}}^j, \mathcal{D}}$, as shown in (11). After the first iteration, the channel input $L_{x_i^j, \mathcal{D}}$ is replaced by $L_{x_i^j, dec} - L_{x_i^j, \mathcal{D}, ext}$ to update $L_{x_i^j, \mathcal{D}, ext}$, i.e.,

$$L_{x_i^j, \mathcal{D}, ext} = 2 \tanh^{-1} \left(\tanh \left(\frac{L_{x_i^j, dec} - L_{x_i^j, \mathcal{D}, ext}}{2} \right) \tanh \left(\frac{L_{x_{\mathcal{R}}^j, \mathcal{D}}}{2} \right) \right). \quad (32)$$

Since the output of the decoder is updated in each iteration, $L_{x_i^j, \mathcal{D}, ext}$ is updated accordingly.

A. Threshold Optimization

Recall that in the uncoded system, the LLR value of $x_{\mathcal{R}}^j$ at the relay, i.e., $L_{x_{\mathcal{R}}^j, \mathcal{R}}$, is approximated as a Gaussian

variable, as shown in (8). In the coded system, since the relay implements network coding based on the decoded LDPC codewords, it is difficult to track the exact PDF of $L_{x_{\mathcal{R}}^j, \mathcal{R}}$. Here, we can still view $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ as a Gaussian variable, and we assume that its variance is twice the absolute value of its mean. This Gaussian approximation is proved in [32] to be accurate enough for a coded system with LDPC codes.

Now, we minimize the BER for the coded system by optimizing the LLR threshold L_T . Recall that in the uncoded system, the values of $\mathbb{E}(P_{i, mrc})$, $\mathbb{E}(P_{i, neg})$, and $\mathbb{E}(P_{i, dir})$, as shown in (23), (25), and (26), respectively, are independent on the threshold L_T . Therefore, we can obtain the closed form of the optimal threshold. On the other hand, in the coded system, the values of $\mathbb{E}(P_{i, mrc})$, $\mathbb{E}(P_{i, neg})$, and $\mathbb{E}(P_{i, dir})$ at the output of the decoders are correlated with the threshold L_T , which complicates the derivation of the optimal threshold. Also, there are no closed forms for these three error probability values in the coded system with LDPC codes. Hence, we focus on how to obtain a sub-optimal threshold in a simple closed form.

One sub-optimal method is that we only minimize the average BER at the input of the decoders (or ‘‘input BER’’ for simplicity) in the first iteration, rather than at the output of the decoders (or ‘‘output BER’’ for simplicity) in the final iteration. This is based on the fact that a good input performance will lead to a good output performance. Since the decoding process has not been performed yet, the calculations of $\mathbb{E}(P_{i, mrc})$, $\mathbb{E}(P_{i, neg})$, and $\mathbb{E}(P_{i, dir})$ at the input of the decoders in the first iteration are the same as those in the uncoded system. The following theorem gives the optimal threshold based on the minimization of the input BER in the first iteration.

Theorem 3: When the statistical CSI of $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{R}\mathcal{D}}$ is available at the relay, the optimal threshold that minimizes the input BER at the destination can be written as

$$L_{T,c}^* = \ln \frac{\mathbb{E}(P_{neg}) - \mathbb{E}(P_{dir})}{\mathbb{E}(P_{dir}) - \mathbb{E}(P_{mrc})}, \quad (33)$$

where $\mathbb{E}(P_{mrc}) = \frac{1}{2}(\mathbb{E}(P_{1, mrc}) + \mathbb{E}(P_{2, mrc}))$, $\mathbb{E}(P_{neg}) = \frac{1}{2}(\mathbb{E}(P_{1, neg}) + \mathbb{E}(P_{2, neg}))$, and $\mathbb{E}(P_{dir}) = \frac{1}{2}(\mathbb{E}(P_{1, dir}) + \mathbb{E}(P_{2, dir}))$. The expressions of $\mathbb{E}(P_{i, mrc})$, $\mathbb{E}(P_{i, neg})$, and $\mathbb{E}(P_{i, dir})$ are given in (23), (25), and (26), respectively.

Proof: We can obtain the proof by following the same method in the proof of *Theorem 1*. ■

We can see that the threshold given in *Theorem 3* has a similar form as that in the uncoded system given in *Theorem 1*. This is due to the fact that similar to the uncoded system, the LLR $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ in the coded system can still be approximated as a Gaussian variable with its variance being twice the absolute value of its mean.

B. BER Evolution

Once the threshold is fixed, we can further investigate the evolution of the BER $P_{e,i}$ in the iterative receiver at the destination. More precisely, we track the statistical CSI based BER $\mathbb{E}(P_{e,i})$ by averaging $P_{e,i}$ over the channels $h_{1\mathcal{D}}$, $h_{2\mathcal{D}}$, and $h_{\mathcal{R}\mathcal{D}}$. Here, by using the term BER evolution, we mean that the BER result is derived in an iterative manner, which mimics the iterative process at the receiver. We first track the BER at the input of the decoder, and then track the BER at the output of the decoder.

Before tracking the BER evolution, we have the following preliminaries. Given the channel parameter $h_{i\mathcal{D}}$, the LLR $L_{x_i^j, \mathcal{D}}$ is a Gaussian variable with mean $m_{L_{i\mathcal{D}}} = \frac{2h_{i\mathcal{D}}^2 x_i^j}{\sigma^2}$. By averaging the PDF of $L_{x_i^j, \mathcal{D}}$ over the channel $h_{i\mathcal{D}}$, we have the expectation of the PDF as [33, 35]

$$\mathbb{E}(p_{L_{x_i^j, \mathcal{D}}}(L)) = \frac{\sigma^2}{2\sqrt{\lambda_{i\mathcal{D}}^2 + 2\lambda_{i\mathcal{D}}\sigma^2}} \exp\left(\frac{x_i^j L - |L| \sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}}}{2}\right). \quad (34)$$

The average channel crossover probability over the channel $h_{i\mathcal{D}}$ is shown in (26) as $\mathbb{E}(P_{i,dir})$.

Similarly to the calculation of $\mathbb{E}(P_{i,dir})$, the crossover probability of the channel $h_{\mathcal{R}\mathcal{D}}$, denoted by $P_{\mathcal{R},dir}$, when averaged over all the possible $h_{\mathcal{R}\mathcal{D}}$, can be written as

$$\mathbb{E}(P_{\mathcal{R},dir}) = \frac{1}{2} \left(1 - \sqrt{\frac{1}{1 + 4\lambda_{\mathcal{R}}}}\right), \quad (35)$$

where

$$\lambda_{\mathcal{R}} = \frac{1}{\frac{2\alpha v_i^2}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}}. \quad (36)$$

The error probability of the network coded symbol $x_{\mathcal{R}}^j$ at the destination has different values corresponding to the three events \mathcal{E}_c , \mathcal{E}_e , and \mathcal{E}_s defined in (6). Specifically, when \mathcal{E}_c happens to the symbol $x_{\mathcal{R}}^j$, it means that the relay has made a correct decision on $x_{\mathcal{R}}^j$. Therefore, the error probability of $x_{\mathcal{R}}^j$ at the destination is exactly the crossover probability of the channel $h_{\mathcal{R}\mathcal{D}}$, i.e., $\mathbb{E}(P_{\mathcal{R},dir})$. When \mathcal{E}_e happens to the symbol $x_{\mathcal{R}}^j$, it means that the relay has made a wrong decision on $x_{\mathcal{R}}^j$, and we calculate the error probability of $x_{\mathcal{R}}^j$ as $1 - \mathbb{E}(P_{\mathcal{R},dir})$. When \mathcal{E}_s happens to the symbol $x_{\mathcal{R}}^j$, we have $v_j = 0$, and the error probability of $x_{\mathcal{R}}^j$ is 0.5.

Based on the above preliminaries, we now track the BER evolution as follows.

1) *BER at the Input of Decoders*: Note that the BER at the input of \mathcal{S}_i 's decoder is equivalent to the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, comb}$. In the sequel, the superscript $[k]$ is attached to variables, e.g., LLRs, to represent these variables in the k -th iteration of the receiver.

To obtain the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, comb}$, we begin with the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, ext}$ in each iteration. In the first iteration, the LLR $L_{x_i^j, \mathcal{D}, ext}^{[1]}$ is calculated based on the received LLRs $L_{x_i^j, \mathcal{D}}$ and $L_{x_i^j, \mathcal{D}}$, as shown in (11). Therefore, the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, ext}$ can be obtained from the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}}$, and the error probability of $x_{\mathcal{R}}^j$ associated with $L_{x_{\mathcal{R}}^j, \mathcal{D}}$. Specifically, when \mathcal{E}_c happens to the symbol $x_{\mathcal{R}}^j$, we have the error probability of x_i^j , denoted by P_{i,ext,\mathcal{E}_c} , as

$$P_{i,ext,\mathcal{E}_c}^{[1]} = \frac{1}{2} (1 - (1 - 2\mathbb{E}(P_{\mathcal{R},dir})) (1 - 2\mathbb{E}(P_{i,dir}))) \\ = \mathbb{E}(P_{\mathcal{R},dir}) + \mathbb{E}(P_{i,dir}) - 2\mathbb{E}(P_{\mathcal{R},dir})\mathbb{E}(P_{i,dir}). \quad (37)$$

When \mathcal{E}_e happens to the symbol $x_{\mathcal{R}}^j$, the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, ext}$, denoted by P_{i,ext,\mathcal{E}_e} , can be calculated as

$$P_{i,ext,\mathcal{E}_e}^{[1]} = \frac{1}{2} (1 - (1 - 2(1 - \mathbb{E}(P_{\mathcal{R},dir}))) (1 - 2\mathbb{E}(P_{i,dir}))) \\ = 1 - P_{i,ext,\mathcal{E}_c}^{[1]}. \quad (38)$$

We now focus on $P_{i,ext,\mathcal{E}_c}^{[k]}$ and $P_{i,ext,\mathcal{E}_e}^{[k]}$, $k > 1$, in the k -th iteration. Note that $L_{x_i^j, \mathcal{D}, ext}^{[k]}$ in the k -th iteration is updated based on the LLR $L_{x_i^j, dec}^{[k-1]} - L_{x_i^j, \mathcal{D}, ext}^{[k-1]}$ from the $(k-1)$ -th iteration. The calculation of $L_{x_i^j, \mathcal{D}, ext}^{[k]}$ can be found in (32).

We denote $P_{i,dec,ext}$ as the error probability of x_i^j associated with the LLR $L_{x_i^j, dec} - L_{x_i^j, \mathcal{D}, ext}$. Note that the calculation of $P_{i,dec,ext}$ is related to the output of \mathcal{S}_i 's decoder, i.e., $L_{x_i^j, dec}$. In the k -th iteration, we can obtain the error probabilities $P_{i,ext,\mathcal{E}_c}^{[k]}$ and $P_{i,ext,\mathcal{E}_e}^{[k]}$ by replacing $\mathbb{E}(P_{i,dir})$ in (37) and (38) with $P_{i,dec,ext}^{[k-1]}$, respectively.

In addition, when \mathcal{E}_s happens to the symbol $x_{\mathcal{R}}^j$, we can see that $\mathbb{E}(P_{\mathcal{R},dir})$ is fixed to a constant 0.5 for all the iterations. In this case, the error probability of x_i^j associated with $L_{x_i^j, \mathcal{D}, ext}$, denoted by P_{i,ext,\mathcal{E}_s} , equals 0.5 for all the iterations.

The LLR $L_{x_i^j, \mathcal{D}, ext}$ in each iteration is approximated as a conditional Gaussian variable given one of the three events \mathcal{E}_c , \mathcal{E}_e , and \mathcal{E}_s , where its variance is assumed to be twice the absolute value of its mean. Corresponding to these three events, the mean of $L_{x_i^j, \mathcal{D}, ext}$ has three possible values, denoted by $m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}$, $m_{L_{i\mathcal{D}, ext, \mathcal{E}_e}}$, and $m_{L_{i\mathcal{D}, ext, \mathcal{E}_s}}$, respectively. More precisely, the mean of $L_{x_i^j, \mathcal{D}, ext}$ equals $m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}$, $m_{L_{i\mathcal{D}, ext, \mathcal{E}_e}}$, and $m_{L_{i\mathcal{D}, ext, \mathcal{E}_s}}$ with probabilities $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$, respectively. In the k -th iteration, these three mean values can be calculated as

$$m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}^{[k]}} = 2x_i^j \left(Q^{-1} \left(P_{i,ext,\mathcal{E}_c}^{[k]} \right) \right)^2, \\ m_{L_{i\mathcal{D}, ext, \mathcal{E}_e}^{[k]}} = -m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}^{[k]}}, \\ m_{L_{i\mathcal{D}, ext, \mathcal{E}_s}^{[k]}} = 0. \quad (39)$$

Note that in (39), the second equation comes from the result $P_{i,ext,\mathcal{E}_e}^{[k]} = 1 - P_{i,ext,\mathcal{E}_c}^{[k]}$, and the third equation holds because of the fact that $P_{i,ext,\mathcal{E}_s}^{[k]} = 0.5$.

With above discussions, we now investigate the BER at the input of \mathcal{S}_i 's decoder, which is equivalent to the error probability of x_i^j associated with the LLR $L_{x_i^j, \mathcal{D}, comb}$. We first derive the PDF of $L_{x_i^j, \mathcal{D}, comb}$, denoted by $p_{L_{x_i^j, \mathcal{D}, comb}}(L)$. Then we obtain the error probability based on this PDF. Since we have $L_{x_i^j, \mathcal{D}, comb} = L_{x_i^j, \mathcal{D}} + L_{x_i^j, \mathcal{D}, ext}$, we can derive the PDF of $L_{x_i^j, \mathcal{D}, comb}$ through convolution of $\mathbb{E}(p_{L_{x_i^j, \mathcal{D}}}(L))$ and the PDF of $L_{x_i^j, \mathcal{D}, ext}$ [33, 35]. Note that $L_{x_i^j, \mathcal{D}, ext}$ has three possible mean values associated with the three events \mathcal{E}_c , \mathcal{E}_e , and \mathcal{E}_s , as shown in (39). For a given mean value of $L_{x_i^j, \mathcal{D}, ext}$, we can derive the corresponding PDF of $L_{x_i^j, \mathcal{D}, comb}$.

Specifically, when the mean value of $L_{x_i^j, \mathcal{D}, ext}$ equals $m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}$, we denote $\sigma_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}^2$ as the variance of $L_{x_i^j, \mathcal{D}, ext}$, and we have $\sigma_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}^2 = 2|m_{L_{i\mathcal{D}, ext, \mathcal{E}_c}}|$. In this

$$\begin{aligned}
p_{L_{x_i^j, \mathcal{D}, \text{comb}}^{[k]}}(L) &= \frac{\sigma^2}{2\sqrt{\lambda_{i\mathcal{D}}^2 + 2\lambda_{i\mathcal{D}}\sigma^2}} \exp\left(-\frac{x_i^j m_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2} + \frac{1 + \frac{\sigma^2}{\lambda_{i\mathcal{D}}}}{4} \left(\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}\right)^2\right) \times \\
&\left[\exp\left(\frac{x_i^j L + \sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} L}{2} + \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} x_i^j \left(\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}\right)^2}{4} - \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} m_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2}\right) \times \right. \\
&Q\left(\frac{L}{\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}} + \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} \sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2} + \frac{x_i^j \sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2} - \frac{m_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}\right) + \\
&\exp\left(\frac{x_i^j L - \sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} L}{2} - \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} x_i^j \left(\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}\right)^2}{4} + \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} m_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2}\right) \times \\
&Q\left(\frac{-L}{\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}} + \frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{i\mathcal{D}}}} \sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2} - \frac{x_i^j \sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{2} + \frac{m_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}{\sigma_{L_{i\mathcal{D}, \text{ext}, \mathcal{E}_c}^{[k]}}}\right) \left. \right]. \tag{40}
\end{aligned}$$

case, the PDF of $L_{x_i^j, \mathcal{D}, \text{comb}}$ in the k -th iteration can be calculated as (40).

The error probability of x_i^j associated with the LLR $L_{x_i^j, \mathcal{D}, \text{comb}}$, denoted by $P_{i, \text{comb}, \mathcal{E}_c}$, can be calculated based on $p_{L_{x_i^j, \mathcal{D}, \text{comb}}}(L)$ in (40). Considering $x_i^j = 1$, we obtain $P_{i, \text{comb}, \mathcal{E}_c}$ by integration

$$P_{i, \text{comb}, \mathcal{E}_c} = \int_{-\infty}^0 p_{L_{x_i^j, \mathcal{D}, \text{comb}}}(L) dL. \tag{41}$$

Then in the k -th iteration, we have $P_{i, \text{comb}, \mathcal{E}_c}^{[k]}$ shown in (42).

We denote $P_{i, \text{comb}, \mathcal{E}_e}$ and $P_{i, \text{comb}, \mathcal{E}_s}$ as the error probabilities of x_i^j associated with $L_{x_i^j, \mathcal{D}, \text{comb}}$ when \mathcal{E}_e and \mathcal{E}_s happen, respectively. In the k -th iteration, we can obtain $P_{i, \text{comb}, \mathcal{E}_e}^{[k]}$ and $P_{i, \text{comb}, \mathcal{E}_s}^{[k]}$ by following the similar calculation in (42).

We view \mathbf{x}_i as an LDPC codeword of code length l . In the k -th iteration of the receiver, at the input of S_i 's decoder, there are average $l\Pr(\mathcal{E}_c)$ symbols in \mathbf{x}_i with BER $P_{i, \text{comb}, \mathcal{E}_c}^{[k]}$, $l\Pr(\mathcal{E}_e)$ symbols in \mathbf{x}_i with BER $P_{i, \text{comb}, \mathcal{E}_e}^{[k]}$, and $l\Pr(\mathcal{E}_s)$ symbols in \mathbf{x}_i with BER $P_{i, \text{comb}, \mathcal{E}_s}^{[k]}$. With these input BERs, we will analyze the BER evolution inside the decoder.

2) *BER at the Output of Decoders*: Assume the sum-product iterative decoding algorithm [29, 30] is applied to each decoder. The superscript $[t, k]$ is attached to variables, e.g., LLRs, to represent these variables in the t -th iteration of the decoder and the k -th iteration of the receiver.

An LDPC code can be defined by a Tanner graph [29, 30] formed by variable nodes and check nodes connected by edges. Assuming d_v (resp. d_c) is the maximum variable (resp. check) node degree, we denote $\lambda(z) = \sum_d \lambda_d z^{d-1}$ (resp. $\rho(z) = \sum_d \rho_d z^{d-1}$) as the variable (resp. check) node degree distribution polynomial of the graph from node perspective. More precisely, λ_d (resp. ρ_d) represents the fraction of variable (resp. check) nodes of degree d . Please refer to [29, 30] for more details about LDPC codes and the sum-product iterative decoding algorithm.

We denote \mathbb{C}_i as the LDPC code implemented at S_i , and denote \mathcal{G}_i as the Tanner graph associated with the code \mathbb{C}_i .

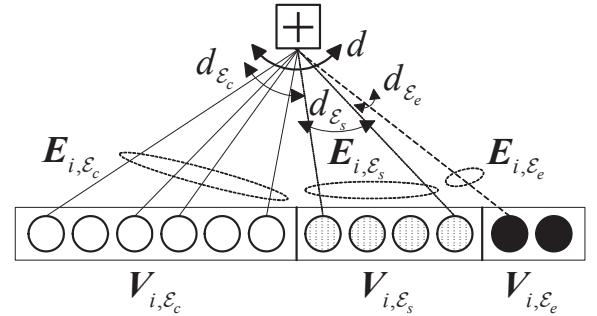


Fig. 3. A subgraph of \mathcal{G}_i , which contains one check node and multiple variable nodes. The variable nodes are divided into three sets, i.e., $\mathbf{V}_{i, \mathcal{E}_c}$, $\mathbf{V}_{i, \mathcal{E}_e}$, and $\mathbf{V}_{i, \mathcal{E}_s}$. The edges are classified into three edge types, i.e., $\mathbf{E}_{i, \mathcal{E}_c}$, $\mathbf{E}_{i, \mathcal{E}_e}$, and $\mathbf{E}_{i, \mathcal{E}_s}$. The degree of the check node is d , in which, $d_{\mathcal{E}_c}$ edges are from the edge type $\mathbf{E}_{i, \mathcal{E}_c}$, $d_{\mathcal{E}_e}$ edges are from the edge type $\mathbf{E}_{i, \mathcal{E}_e}$, and $d_{\mathcal{E}_s}$ edges are from the edge type $\mathbf{E}_{i, \mathcal{E}_s}$.

Also, we divide all the variable nodes in \mathcal{G}_i into three sets, namely, $\mathbf{V}_{i, \mathcal{E}_c}$, $\mathbf{V}_{i, \mathcal{E}_e}$, and $\mathbf{V}_{i, \mathcal{E}_s}$, which are associated with the symbols in the codeword \mathbf{x}_i whose input BERs are $P_{i, \text{comb}, \mathcal{E}_c}$, $P_{i, \text{comb}, \mathcal{E}_e}$, and $P_{i, \text{comb}, \mathcal{E}_s}$, respectively. Accordingly, all the edges in \mathcal{G}_i are classified into three edge types, namely, $\mathbf{E}_{i, \mathcal{E}_c}$, $\mathbf{E}_{i, \mathcal{E}_e}$, and $\mathbf{E}_{i, \mathcal{E}_s}$. That is, edges emanating from the variable nodes in the sets $\mathbf{V}_{i, \mathcal{E}_c}$, $\mathbf{V}_{i, \mathcal{E}_e}$, and $\mathbf{V}_{i, \mathcal{E}_s}$ are classified into the types $\mathbf{E}_{i, \mathcal{E}_c}$, $\mathbf{E}_{i, \mathcal{E}_e}$, and $\mathbf{E}_{i, \mathcal{E}_s}$, respectively. It is clear that an edge belongs to the edge types $\mathbf{E}_{i, \mathcal{E}_c}$, $\mathbf{E}_{i, \mathcal{E}_e}$, and $\mathbf{E}_{i, \mathcal{E}_s}$ with probabilities $\Pr(\mathcal{E}_c)$, $\Pr(\mathcal{E}_e)$, and $\Pr(\mathcal{E}_s)$, respectively.

Fig. 3 shows a subgraph of \mathcal{G}_i with a check node of degree d and multiple variable nodes. The three sets of variable nodes and the three edge types are illustrated in Fig. 3. Also, we can see that the check node in Fig. 3 have $d_{\mathcal{E}_c}$, $d_{\mathcal{E}_e}$, and $d_{\mathcal{E}_s}$ edges in the edge types $\mathbf{E}_{i, \mathcal{E}_c}$, $\mathbf{E}_{i, \mathcal{E}_e}$, and $\mathbf{E}_{i, \mathcal{E}_s}$, respectively, where $d = d_{\mathcal{E}_c} + d_{\mathcal{E}_e} + d_{\mathcal{E}_s}$. We denote $\mathbb{C}_{i, (d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})}$ as the set of check nodes which have the same degree property as the one in Fig. 3. For all the check nodes of degree d , we denote $\zeta(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$ as the fraction of the check nodes that belong to $\mathbb{C}_{i, (d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})}$. From the expansion

$$P_{i,comb,\mathcal{E}_c}^{[k]} = Q \left(\frac{m_{L_{iD},ext,\mathcal{E}_c}^{[k]}}{\sigma_{L_{iD},ext,\mathcal{E}_c}^{[k]}} \right) + \sigma^2 \exp \left(-\frac{m_{L_{iD},ext,\mathcal{E}_c}^{[k]}}{2} + \frac{1 + \frac{\sigma^2}{\lambda_{iD}}}{4} \left(\sigma_{L_{iD},ext,\mathcal{E}_c}^{[k]} \right)^2 \right) \times$$

$$\left(\frac{\exp \left(\frac{\sqrt{1 + \frac{2\sigma^2}{\lambda_{iD}}} \left(\sigma_{L_{iD},ext,\mathcal{E}_c}^{[k]} \right)^2 - 2m_{L_{iD},ext,\mathcal{E}_c}^{[k]}}{4} \right)}{\sqrt{\lambda_{iD}^2 + 2\lambda_{iD}\sigma^2} \left(1 + \sqrt{1 + \frac{2\sigma^2}{\lambda_{iD}}} \right)} Q \left(\frac{(1 + \sqrt{1 + \frac{2\sigma^2}{\lambda_{iD}}}) \sigma_{L_{iD},ext,\mathcal{E}_c}^{[k]} - \frac{m_{L_{iD},ext,\mathcal{E}_c}^{[k]}}{\sigma_{L_{iD},ext,\mathcal{E}_c}^{[k]}}}{2}} \right)}{\sqrt{\lambda_{iD}^2 + 2\lambda_{iD}\sigma^2} \left(\sqrt{1 + \frac{2\sigma^2}{\lambda_{iD}}} - 1 \right)} \right). \quad (42)$$

of $(\Pr(\mathcal{E}_c) + \Pr(\mathcal{E}_e) + \Pr(\mathcal{E}_s))^d$, we can obtain that

$$\zeta(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s}) = \frac{d!}{d_{\mathcal{E}_c}! d_{\mathcal{E}_e}! d_{\mathcal{E}_s}!} (\Pr(\mathcal{E}_c))^{d_{\mathcal{E}_c}} (\Pr(\mathcal{E}_e))^{d_{\mathcal{E}_e}} (\Pr(\mathcal{E}_s))^{d_{\mathcal{E}_s}}. \quad (43)$$

Within the edge type $\mathbf{E}_{i,\mathcal{E}_c}$ (resp. $\mathbf{E}_{i,\mathcal{E}_e}$, $\mathbf{E}_{i,\mathcal{E}_s}$), we denote $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_c}}$ (resp. $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_e}}$, $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_s}}$) as the fraction of the edges connected to the check nodes in the set $\mathbf{C}_{i,(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}$. As the total number of edges in $\mathbf{E}_{i,\mathcal{E}_c}$ after normalization is $\Pr(\mathcal{E}_c) \sum_{d_1=0}^{d_c} d_1 \rho_{d_1}$, we have

$$\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_c}} = \frac{d_{\mathcal{E}_c} \zeta(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s}) \rho_d}{\Pr(\mathcal{E}_c) \sum_{d_1=0}^{d_c} d_1 \rho_{d_1}}. \quad (44)$$

We can obtain $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_e}}$ (resp. $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_s}}$) by following the same method as in (44).

In the t -th iteration of the decoder and the k -th iteration of the receiver, we denote $P_{i,var,\mathcal{E}_c}^{[t,k]}$ (resp. $P_{i,var,\mathcal{E}_e}^{[t,k]}$, $P_{i,var,\mathcal{E}_s}^{[t,k]}$) as the average BER associated with the LLRs sent from the variable nodes in $\mathbf{V}_{i,\mathcal{E}_c}$ (resp. $\mathbf{V}_{i,\mathcal{E}_e}$, $\mathbf{V}_{i,\mathcal{E}_s}$) to the check nodes, on each edge of the edge type $\mathbf{E}_{i,\mathcal{E}_c}$ (resp. $\mathbf{E}_{i,\mathcal{E}_e}$, $\mathbf{E}_{i,\mathcal{E}_s}$). Also, we denote $P_{i,chk,\mathcal{E}_c}^{[t,k]}$ (resp. $P_{i,chk,\mathcal{E}_e}^{[t,k]}$, $P_{i,chk,\mathcal{E}_s}^{[t,k]}$) as the average BER associated with the LLRs sent from the check nodes to the variable nodes in $\mathbf{V}_{i,\mathcal{E}_c}$ (resp. $\mathbf{V}_{i,\mathcal{E}_e}$, $\mathbf{V}_{i,\mathcal{E}_s}$), on each edge of the edge type $\mathbf{E}_{i,\mathcal{E}_c}$ (resp. $\mathbf{E}_{i,\mathcal{E}_e}$, $\mathbf{E}_{i,\mathcal{E}_s}$). Initially, we have $P_{i,var,\mathcal{E}_c}^{[1,k]} = P_{i,comb,\mathcal{E}_c}^{[1,k]}$, $P_{i,var,\mathcal{E}_e}^{[1,k]} = P_{i,comb,\mathcal{E}_e}^{[1,k]}$, and $P_{i,var,\mathcal{E}_s}^{[1,k]} = P_{i,comb,\mathcal{E}_s}^{[1,k]}$.

Now, we investigate the BER evolution from the check nodes to the variable nodes in the t -th iteration of the decoder and the k -th iteration of the receiver. Without a loss of generality, we focus on the derivation of $P_{i,chk,\mathcal{E}_c}^{[t,k]}$. Note that $P_{i,chk,\mathcal{E}_c}^{[t,k]}$ is an average BER associated with the LLRs within the edge type $\mathbf{E}_{i,\mathcal{E}_c}$. For a given vector $(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$, we denote $P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$ as the BER associated with the LLRs sent from the check nodes in $\mathbf{C}_{i,(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}$ to

the variable nodes in $\mathbf{V}_{i,\mathcal{E}_c}$. Then we have

$$P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s}) = \frac{1}{2} \left(1 - \left(1 - 2P_{i,var,\mathcal{E}_c}^{[t,k]} \right)^{d_{\mathcal{E}_c-1}} \left(1 - 2P_{i,var,\mathcal{E}_e}^{[t,k]} \right)^{d_{\mathcal{E}_e}} \left(1 - 2P_{i,var,\mathcal{E}_s}^{[t,k]} \right)^{d_{\mathcal{E}_s}} \right). \quad (45)$$

In (45), if $d_{\mathcal{E}_c} = 0$, we set $P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$ to zero. Since $P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$ in (45) varies according to different $(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$, we average $P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s})$ over all possible values of $d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}$, and $d_{\mathcal{E}_s}$. We hence obtain the average BER $P_{i,chk,\mathcal{E}_c}^{[t,k]}$ as

$$P_{i,chk,\mathcal{E}_c}^{[t,k]} = \sum_{d=0}^{d_c} \sum_{d_{\mathcal{E}_c}=0}^d \sum_{d_{\mathcal{E}_e}=0}^{d-d_{\mathcal{E}_c}} P_{i,chk,\mathcal{E}_c}^{[t,k]}(d, d_{\mathcal{E}_c}, d_{\mathcal{E}_e}, d_{\mathcal{E}_s}) \lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_c}}. \quad (46)$$

Similarly, we can obtain $P_{i,chk,\mathcal{E}_e}^{[t,k]}$ (resp. $P_{i,chk,\mathcal{E}_s}^{[t,k]}$) by using $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_e}}$ (resp. $\lambda_{(d,d_{\mathcal{E}_c},d_{\mathcal{E}_e},d_{\mathcal{E}_s})}^{d_{\mathcal{E}_s}}$) and following the same method as in (46).

Within each edge type, we make Gaussian approximations on the LLRs sent from the check nodes to the variable nodes, where the variances of these LLRs are twice the absolute values of their means. We denote m_{i,chk,\mathcal{E}_c} (resp. m_{i,chk,\mathcal{E}_e} , m_{i,chk,\mathcal{E}_s}) as the mean, and denote $\sigma_{i,chk,\mathcal{E}_c}^2$ (resp. $\sigma_{i,chk,\mathcal{E}_e}^2$, $\sigma_{i,chk,\mathcal{E}_s}^2$) as the variance, of the LLR sent from the check nodes to each variable node in the set $\mathbf{V}_{i,\mathcal{E}_c}$ (resp. $\mathbf{V}_{i,\mathcal{E}_e}$, $\mathbf{V}_{i,\mathcal{E}_s}$). Then we have $|m_{i,chk,\mathcal{E}_c}^{[t,k]}| = 2 \left(Q^{-1} \left(P_{i,chk,\mathcal{E}_c}^{[t,k]} \right) \right)^2$, $|m_{i,chk,\mathcal{E}_e}^{[t,k]}| = 2 \left(Q^{-1} \left(P_{i,chk,\mathcal{E}_e}^{[t,k]} \right) \right)^2$, and $|m_{i,chk,\mathcal{E}_s}^{[t,k]}| = 2 \left(Q^{-1} \left(P_{i,chk,\mathcal{E}_s}^{[t,k]} \right) \right)^2$. Also, we have $\left(\sigma_{i,chk,\mathcal{E}_c}^{[t,k]} \right)^2 = 2 |m_{i,chk,\mathcal{E}_c}^{[t,k]}|$, $\left(\sigma_{i,chk,\mathcal{E}_e}^{[t,k]} \right)^2 = 2 |m_{i,chk,\mathcal{E}_e}^{[t,k]}|$, and $\left(\sigma_{i,chk,\mathcal{E}_s}^{[t,k]} \right)^2 = 2 |m_{i,chk,\mathcal{E}_s}^{[t,k]}|$.

Next, we investigate the BER evolution from the variable nodes to the check nodes in the $(t+1)$ -th iteration of the decoder and the k -th iteration of the receiver. Without a loss of generality, we focus on the variable node of degree d

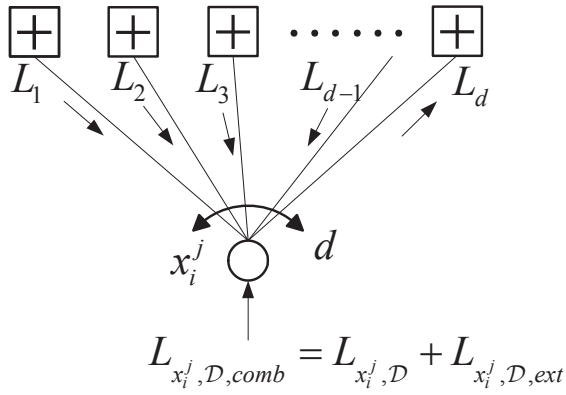


Fig. 4. The variable node of degree d associated with the symbol x_i^j . The value of the LLR L_d sent from this variable node to a check node is calculated based on the input LLR $L_{x_i^j, \mathcal{D}, comb}$ and the LLRs L_1, L_2, \dots, L_{d-1} sent from other $d-1$ check nodes to this variable node.

associated with the symbol x_i^j . Fig. 4 shows the message transfer. The calculation of L_d is based on the input LLR $L_{x_i^j, \mathcal{D}, comb}$ and the LLRs L_1, L_2, \dots, L_{d-1} from other $d-1$ check nodes. That is

$$L_d = L_{x_i^j, \mathcal{D}, comb} + \sum_{\iota=1}^{d-1} L_{d_\iota} = L_{x_i^j, \mathcal{D}} + L_{x_i^j, \mathcal{D}, ext} + \sum_{\iota=1}^{d-1} L_{d_\iota}. \quad (47)$$

Since $L_{x_i^j, \mathcal{D}, ext}$ and L_ι , $\iota = 1, \dots, d-1$, are approximated as Gaussian variables, we can view $L_{x_i^j, \mathcal{D}, ext} + \sum_{\iota=1}^{d-1} L_{d_\iota}$ as a Gaussian variable. The PDF of L_d can be obtained through the convolution of $\mathbb{E}(p_{L_{x_i^j, \mathcal{D}}}(L))$ and the PDF of $L_{x_i^j, \mathcal{D}, ext} + \sum_{\iota=1}^{d-1} L_{d_\iota}$, based on which, we can obtain the BER associated with L_D . In fact, we can use (42) to obtain the BER associated with L_D by replacing $m_{L_{iD}, ext, \mathcal{E}_c}^{[k]}$ and $(\sigma_{L_{iD}, ext, \mathcal{E}_c}^{[k]})^2$ in (42) with the mean and variance of $L_{x_i^j, \mathcal{D}, ext} + \sum_{\iota=1}^{d-1} L_{d_\iota}$, respectively. Assuming that the variable node associated with x_i^j is in the set $\mathbf{V}_{i, \mathcal{E}_c}$, we define the BER associated with L_d by $P_{i, var, \mathcal{E}_c}^{[t+1, k]}(d)$. The mean and variance of $L_{x_i^j, \mathcal{D}, ext} + \sum_{\iota=1}^{d-1} L_{d_\iota}$ can be calculated as $m_{L_{iD}, ext, \mathcal{E}_c}^{[k]} + (d-1)m_{i, chk, \mathcal{E}_c}^{[t, k]}$ and $(\sigma_{L_{iD}, ext, \mathcal{E}_c}^{[k]})^2 + (d-1)(\sigma_{i, chk, \mathcal{E}_c}^{[t, k]})^2$, respectively. Hence, we can obtain $P_{i, var, \mathcal{E}_c}^{[t+1, k]}(d)$ by using (42) with the corresponding mean and variance. By averaging $P_{i, var, \mathcal{E}_c}^{[t+1, k]}(d)$ over all the variable degrees, we have

$$P_{i, var, \mathcal{E}_c}^{[t+1, k]} = \sum_{d=0}^{d_v} \frac{\lambda_d d}{\sum_{d_1=0}^{d_v} \lambda_{d_1} d_1} P_{i, var, \mathcal{E}_c}^{[t+1, k]}(d). \quad (48)$$

By following the similar method, we can obtain $P_{i, var, \mathcal{E}_e}^{[t+1, k]}$ and $P_{i, var, \mathcal{E}_s}^{[t+1, k]}$. Based on $P_{i, var, \mathcal{E}_c}^{[t+1, k]}$, $P_{i, var, \mathcal{E}_e}^{[t+1, k]}$, and $P_{i, var, \mathcal{E}_s}^{[t+1, k]}$, we can thus update the BERs from the check nodes to the variable nodes in the $t+1$ -th iteration of the decoder, i.e., $P_{i, chk, \mathcal{E}_c}^{[t+1, k]}$, $P_{i, chk, \mathcal{E}_e}^{[t+1, k]}$, and $P_{i, chk, \mathcal{E}_s}^{[t+1, k]}$, by following (45) and (46).

Assume there are total K iterations in the receiver, and in each iteration of the receiver, there are total T iterations of the decoding, we obtain the final BER results at the output of the receiver as $P_{i, var, \mathcal{E}_c}^{[T, K]}$, $P_{i, var, \mathcal{E}_e}^{[T, K]}$, and $P_{i, var, \mathcal{E}_s}^{[T, K]}$. These BER

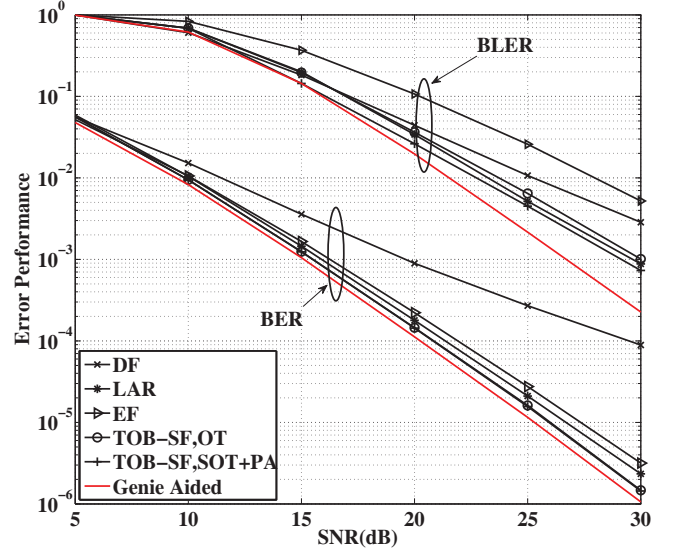


Fig. 5. Error performance for Case 1, i.e., the strong source-to-relay link scenario, in the uncoded system.

results will be used to optimize the power allocation in the following subsection.

C. Power Allocation

We optimize the power allocation α at the relay to further improve the coding gain. Recall that in the uncoded system, we have derived the optimal power allocation α^* in (31) to improve the BER. Although this α^* can be directly applied to the coded system (the probabilities $\Pr(\mathcal{E}_e)$ and $\Pr(\mathcal{E}_c)$ in (31) need to be recalculated in the coded system) to improve the input BER, it cannot guarantee the optimality of the output BER in the coded system.

Here, to optimize the power allocation for the coded system, we track the BER evolution in each iteration, and minimize the output BER at the end of the final iteration. Assume there are total K iterations in the receiver, and in each iteration of the receiver, there are total T iterations of the decoding. We fix the threshold and calculate the statistical CSI based BER $\mathbb{E}(P_{e, i})$ as

$$\mathbb{E}(P_{e, i}) = P_{i, var, \mathcal{E}_c}^{[T, K]} \Pr(\mathcal{E}_e) + P_{i, var, \mathcal{E}_e}^{[T, K]} \Pr(\mathcal{E}_e) + P_{i, var, \mathcal{E}_s}^{[T, K]} \Pr(\mathcal{E}_s). \quad (49)$$

We optimize the power allocation α by searching from 0 to 1, and obtain the optimal power allocation α which leads to the minimum value of $\frac{1}{2}(\mathbb{E}(P_{e, 1}) + \mathbb{E}(P_{e, 2}))$.

VI. SIMULATIONS

In the simulations, we set the frame (or codeword) length as $l = 10,000$ for the uncoded (or coded) system. Hence, each transmission period is of length 30,000. As we consider quasi-static fading channels, the channel coefficients h_{1R} , h_{2R} , h_{1D} , h_{2D} , and h_{RD} are constant in each transmission period, and change independently from one period to another.

We focus on a symmetric scenario where (i) the two sources have the same distance to the relay and the same distance to

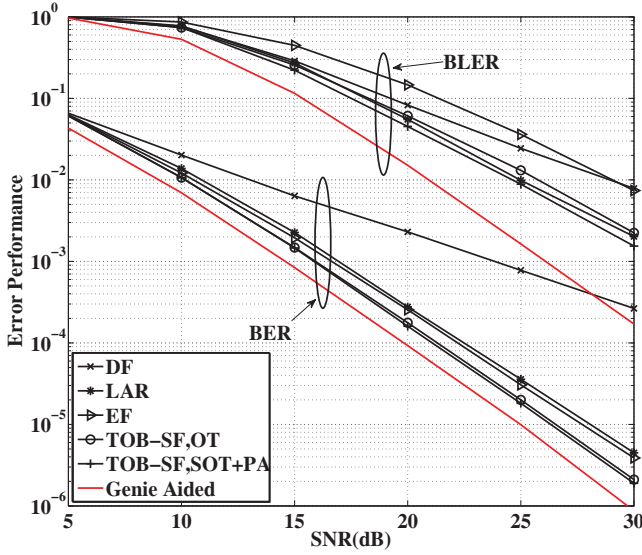


Fig. 6. Error performance for Case 2, i.e., the symmetric channel scenario, in the uncoded system.

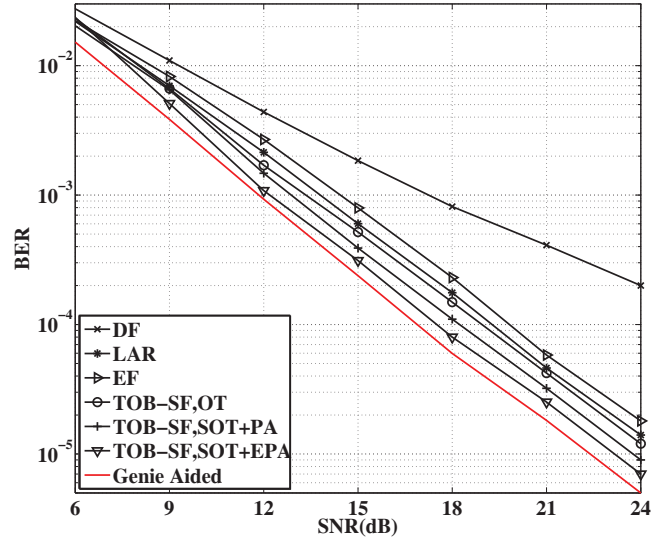


Fig. 8. BER performance for Case 1 in the LDPC coded system.

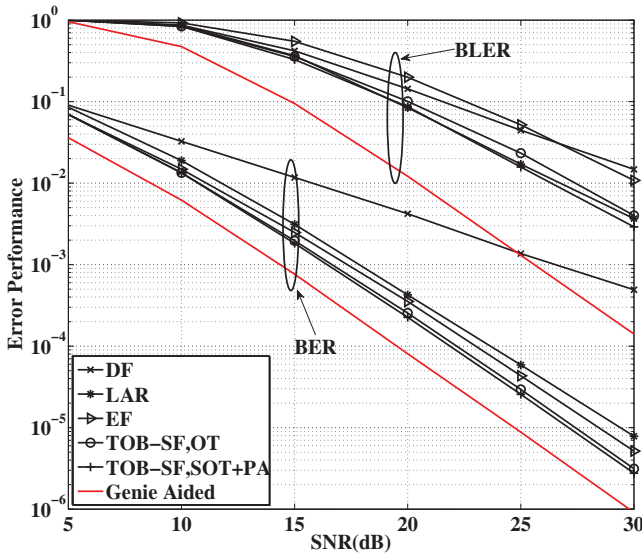


Fig. 7. Error performance for Case 3, i.e., the strong relay-to-destination link scenario, in the uncoded system.

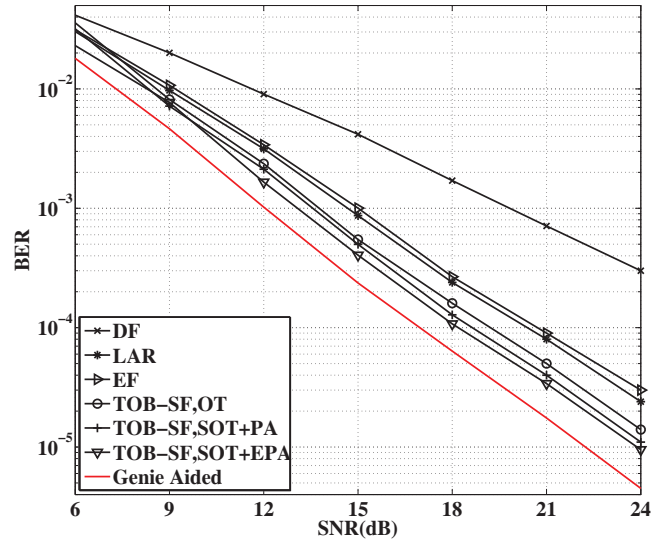


Fig. 9. BER performance for Case 2 in the LDPC coded system.

the destination, and (ii) the two sources, the relay, and the destination are aligned on the same horizontal line. The distances between the sources and the destination are normalized as one, i.e., $d_{1D} = d_{2D} = 1$. The relay is located between the sources and the destination. In our simulations, we consider three cases, namely, Case 1, Case 2, and Case 3, according to the three different locations of the relay. Specifically, In Case 1, we consider a strong source-to-relay link scenario, where $d_{1R} = d_{2R} = 0.3$, and $d_{RD} = 0.7$; In Case 2, we consider the symmetric scenario, where $d_{1R} = d_{2R} = 0.5$, and $d_{RD} = 0.5$; In Case 3, we consider a strong relay-to-destination scenario, where $d_{1R} = d_{2R} = 0.7$, and thus $d_{RD} = 0.3$. Also, we set the attenuation exponent $\gamma = 2$. Hence, we can obtain the

channel gains λ_{1R} , λ_{2R} , λ_{1D} , λ_{2D} , and λ_{RD} based on γ and the corresponding distances.

We now evaluate the BER performance in the uncoded system. Besides the BER, we also investigate the block error rates (BLER) in the uncoded system, where a block consists of two frames from the two sources. Fig. 5 shows the error performance for Case 1. In Fig. 5, the proposed TOB-SF protocol with the optimal threshold (OT) L_T^* is denoted by ‘TOB-SF, OT’, and the proposed TOB-SF protocol with the sub-optimal threshold (SOT) $\ln \rho$ plus the optimal power allocation α^* (shown in (31)) is denoted by ‘TOB-SF, SOT+PA’. We use the conventional DF, network coded EF (i.e., the network-coded soft bit forwarding protocol proposed in [36]), and the link-adaptive regeneration (LAR) protocol (i.e., the power scaling scheme in [27]) as benchmarks, denoted by

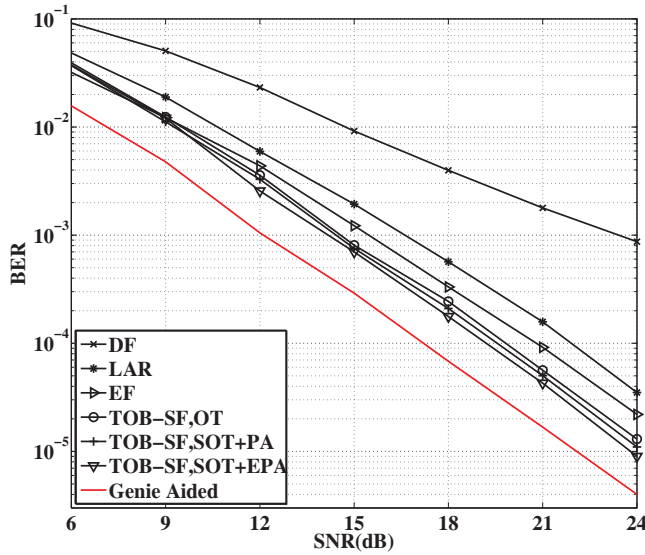


Fig. 10. BER performance for Case 3 in the LDPC coded system.

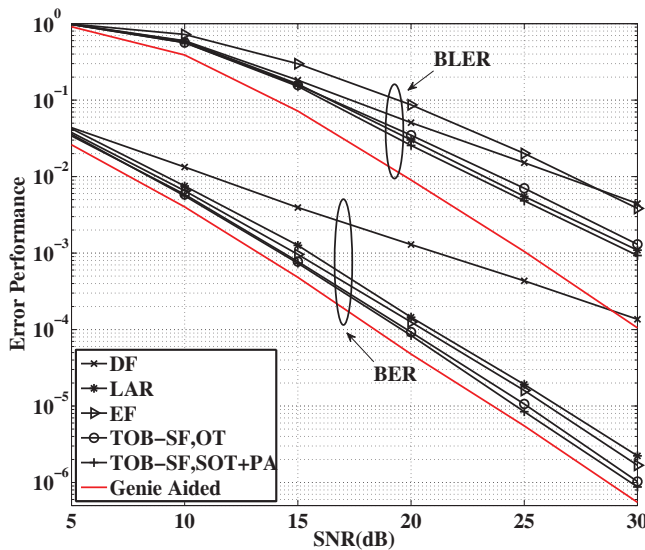


Fig. 11. BER performance for the asymmetric case in the uncoded system.

'DF', 'EF', and 'LAR', respectively. Note that the LAR has a lower computation complexity compared with our TOB-SF protocols. Specifically, in the LAR, one multiplication operation is needed for each network coded symbol at the relay, while in the TOB-SF, the calculation of the LLR value for each network coded symbol is needed, which includes the calculations of two $\tan(\cdot)$ functions and one $\tan^{-1}(\cdot)$ function. Also, 'Genie Aided' protocol in Fig. 5 is a desired protocol, where the source-to-relay channels are error free.

From Fig. 5 we can see that 'DF' can only achieve a diversity of one, while all the other protocols can achieve the full diversity gain of the system, i.e., a diversity of two. Our 'TOB-SF, OT' and 'TOB-SF, SOT+PA' outperform 'DF', 'EF', and 'LAR' in terms of BER performance, and are close to the 'Genie Aided'. Also, 'TOB-SF, SOT+PA' is slightly better than

'TOB-SF, OT'. The reason why the BER performance of the two TOB-SF protocols is better than that of the LAR is that our TOB-SF protocols are optimized to minimize the system BER, while the LAR is an intuitive power control protocol, which cannot guarantee the optimality of the system BER. For the BLER performance, our 'TOB-SF, SOT+PA' outperforms 'DF', 'EF', 'LAR', and 'TOB-SF, OT'. 'LAR' is better than 'TOB-SF, OT'. And 'EF' performs the worst.

Fig. 6 and Fig. 7 show the error performance of Case 2 and Case 3, respectively. In both cases, our 'TOB-SF, OT' and 'TOB-SF, SOT+PA' can achieve the full diversity gain, and outperform the 'DF', 'EF' and 'LAR' in terms of BER performance. Also, 'TOB-SF, SOT+PA' has a better BER performance than 'TOB-SF, OT'. For the BLER performance in the two cases, the 'TOB-SF, SOT+PA' is always better than the 'LAR', while the 'LAR' is always better than 'TOB-SF, OT'. And the 'DF' and 'EF' always perform the worst.

From Fig. 5 to Fig. 7, it is clear that 'TOB-SF, OT' outperforms 'LAR' in terms of BER, since 'TOB-SF, OT' is optimized to minimize the BER, while 'LAR' cannot guarantee the optimality of the BER performance. The reason why 'LAR' outperforms 'TOB-SF, OT' in terms of BLER is because that 'LAR' adapts the power for each transmitted block according to the channel information, while 'TOB-SF, OT' does not implement any power control. Also, we can see that 'TOB-SF, SOT+PA', which optimizes the power at the relay, outperforms 'TOB-SF, OT' and 'LAR' in terms of both the BER and BLER performance. This again implies that power control is needed to improve the BLER performance.

Next, we investigate the error performance for the LDPC coded system. Here, the LDPC codes in both sources are designed with the same code rate, and the same degree profile. Specifically, we fix the code rate as 0.7, and the code profile is as follows. From the variable node point of view, the degree distribution is $\lambda(z) = 0.4183224z + 0.3317611z^2 + 0.069378148z^5 + 0.10453512z^6 + 0.0760031z^{19}$. From the check node point of view, the degree distribution is $\rho(z) = z^{14}$.

Fig. 8, Fig. 9, and Fig. 10 shows the BER performance of Case1, Case2, and Case3, respectively, for the LDPC coded system. Again, we use 'DF', 'EF', and 'LAR' as benchmarks. In the three figures, 'TOB-SF, OT' represents the TOB-SF protocol with the threshold shown in (33), and 'TOB-SF, SOT+PA' represents the TOB-SF protocol with the threshold $\ln \rho$ and the power allocation α^* in (31). Note that both 'TOB-SF, OT' and 'TOB-SF, SOT+PA' are the same as that in the uncoded system. That is, they optimize the input BER rather than the output BER of the receiver at the destination in the coded system, which leads to a sub-optimal BER performance. Also, in the three figures, 'TOB-SF, SOT+EPA' represents the TOB-SF protocol with threshold $\ln \rho$ plus the BER evolution based power allocation.

From the three figures, we can see that the 'DF' can only achieve a diversity of one, while all the other protocols can achieve the full diversity gain. Compared with 'LAR' and 'EF', the proposed three TOB-SF protocols have better BER performance. Similar to the uncoded system, the performance gap between the benchmarks and the proposed TOB-SF protocols become large when the relay is closer to the destination. Interestingly, 'LAR' performs worse and worse when the relay

becomes closer to the destinations. Among the three TOB-SF protocols, ‘TOB-SF, SOT+PA’ outperforms ‘TOB-SF, OT’, while ‘TOB-SF, SOT+EPA’ is always the best. This is because ‘TOB-SF, SOT+EPA’ optimizes the output BER.

Based on the above discussions, the proposed TOB-SF protocols always perform better than the benchmarks ‘DF’, ‘EF’, and ‘LAR’. Although ‘TOB-SF, SOT+EPA’ performs the best, the other two TOB-SF protocols, i.e., ‘TOB-SF, SOT+PA’ and ‘TOB-SF, OT’ are simple and easy to implement with a slight error performance loss.

Furthermore, we consider an asymmetric scenario, where the two sources are asymmetrically located. Specifically, we assume $d_{1\mathcal{R}} = 0.5, d_{2\mathcal{R}} = 0.2, d_{\mathcal{R}\mathcal{D}} = 0.5, d_{1\mathcal{D}} = 1.0,$ and $d_{2\mathcal{D}} = 0.7$. We simulate the BER and BLER performance for the uncoded case in this scenario. The error performance is shown in Fig. 11. From this figure, we can see that our ‘TOB-SF, OT’ and ‘TOB-SF, SOT+PA’ outperform the ‘DF’, ‘EF’ and ‘LAR’ in terms of the BER performance. Also, ‘TOB-SF, SOT+PA’ has a better BER performance than ‘TOB-SF, OT’. For the BLER performance, the ‘TOB-SF, SOT+PA’ is better than the ‘LAR’, and the ‘LAR’ is better than ‘TOB-SF, OT’. The simulation results show that the performance comparisons of these protocols in asymmetric scenarios are consistent with those in the symmetric scenarios.

VII. CONCLUSION

In this paper, we propose and optimize a TOB-SF protocol for a multi-source relay system with network coding. We consider both the uncoded and LDPC coded systems. In the uncoded system, we first derive the BER expressions at the destination, and develop the optimal threshold. Then we prove that the TOB-SF protocol can achieve the full diversity gain by using the optimal threshold. Furthermore, we optimize the power allocation at the relay to achieve a better performance. In the coded system with LDPC codes, we develop a methodology to track the BER evolution by using Gaussian approximations, and optimize the power allocation at the relay based on the BER evolution results. Simulation results show that the proposed TOB-SF protocol with power allocation outperforms other conventional relaying protocols in terms of error performance.

APPENDIX A PROOF OF THEOREM 1

We take the derivation of the average BER $\mathbb{E}(P_e) = \frac{1}{2}(\mathbb{E}(P_{e,1}) + \mathbb{E}(P_{e,2}))$ in terms of L_T and let the result equal zero. Then we obtain

$$\begin{aligned} & (\mathbb{E}(P_{dir}) - \mathbb{E}(P_{mrc})) \frac{\partial \Pr(\mathcal{E}_c)}{\partial L_T} = \\ & (\mathbb{E}(P_{neg}) - \mathbb{E}(P_{dir})) \frac{\partial \Pr(\mathcal{E}_e)}{\partial L_T}. \quad (50) \end{aligned}$$

For the case $x_1^j x_2^j = 1$, $\Pr(\mathcal{E}_c)$ and $\Pr(\mathcal{E}_e)$ are shown in (9). Here, the LLR $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ is approximated as a Gaussian variable with its variance being twice the absolute value of its mean. Based on the PDF of $L_{x_{\mathcal{R}}^j, \mathcal{R}}$ in (8) and the fact that $\sigma_{L_{\oplus}}^2 = 2m_{L_{\oplus}}$, the optimal threshold L_T can be derived as shown in (10). For the case $x_1^j x_2^j = -1$, we can obtain the same result on the optimal L_T . This completes the proof. ■

APPENDIX B PROOF OF LEMMA 1

We define $x \triangleq \frac{2h_{i\mathcal{D}}^2}{\sigma^2}$, $y \triangleq \frac{2h_{\bar{i}\mathcal{D}}^2}{\sigma^2}$, and $v \triangleq \frac{2\alpha h_{\mathcal{R}\mathcal{D}}^2}{\sigma^2}$. The PDFs of the three variables can be written as follows.

$$\begin{aligned} p_x(x) &= \frac{1}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}} \exp\left(-\frac{x}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}}\right), \\ p_y(y) &= \frac{1}{\frac{2}{\sigma^2} \lambda_{\bar{i}\mathcal{D}}} \exp\left(-\frac{y}{\frac{2}{\sigma^2} \lambda_{\bar{i}\mathcal{D}}}\right), \\ p_v(v) &= \frac{1}{\frac{2\alpha}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}} \exp\left(-\frac{v}{\frac{2\alpha}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}}\right). \end{aligned} \quad (51)$$

According to [37], the PDF of $w = \min(y, v)$ can be calculated as

$$p_w(w) = p_y(w)(1 - F_v(w)) + p_v(w)(1 - F_y(w)), \quad (52)$$

where $F_y(w) = \int_{-\infty}^w p_x(u) du$ and $F_v(w) = \int_{-\infty}^w p_v(u) du$. Then we obtain

$$\begin{aligned} p_w(w) &= \\ & \left(\frac{1}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}} + \frac{1}{\frac{2\alpha}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}} \right) \exp\left(-\left(\frac{1}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}} + \frac{1}{\frac{2\alpha}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}}\right) w\right). \end{aligned} \quad (53)$$

To simplify the notations, we replace $\frac{1}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}}$ with λ_i , and replace $\frac{1}{\frac{2}{\sigma^2} \lambda_{i\mathcal{D}}} + \frac{1}{\frac{2\alpha}{\sigma^2} \lambda_{\mathcal{R}\mathcal{D}}}$ with $\bar{\lambda}_i$. By viewing x as $m_{L_{i\mathcal{D}}, dir}$, we can directly obtain the PDF of $m_{L_{i\mathcal{D}}, dir}$ from (51). Next, we derive the PDFs for the two variables $g = x + w$ and $q = \frac{(x-w)^2}{x+w}$. According to [37], we have

$$p_g(g) = \frac{\lambda_i \bar{\lambda}_i}{\lambda_i - \bar{\lambda}_i} (\exp(-\bar{\lambda}_i g) - \exp(-\lambda_i g)). \quad (54)$$

To obtain the PDF of q , we first focus on the PDF of $r = x - w$, and we have

$$p_r(r) = \frac{\lambda_i \bar{\lambda}_i}{\lambda_i + \bar{\lambda}_i} \exp(-\lambda_i r). \quad (55)$$

Since $q = \frac{r^2}{g}$, based on $p_g(g)$ and $p_r(r)$, we have [22, 37]

$$\begin{aligned} p_q(q) &= \frac{\lambda_i^2 \bar{\lambda}_i^2}{\lambda_i^2 - \bar{\lambda}_i^2} \int_0^\infty \frac{u^2}{\sqrt{q}} \\ & (\exp(-\bar{\lambda}_i u^2 - \lambda_i \sqrt{q} u) - \exp(-\lambda_i u^2 - \bar{\lambda}_i \sqrt{q} u)) du. \end{aligned} \quad (56)$$

By viewing g as $m_{L_{i\mathcal{D}}, mrc}$, and viewing q as $m_{L_{i\mathcal{D}}, neg}$, we can obtain the PDFs for the variables $m_{L_{i\mathcal{D}}, mrc}$ and $m_{L_{i\mathcal{D}}, neg}$, respectively. Hence, we complete the proof. ■

APPENDIX C PROOF OF THEOREM 2

Without a loss of generality, we consider the case $x_1^j x_2^j = 1$, and focus on the probabilities in (9). In (8), the mean value $m_{L_{\oplus}}$ can be approximated by following the same method in (16) in the high SNR region, and we have $m_{L_{\oplus}} = \frac{2}{\sigma^2} \min(h_{1\mathcal{R}}^2, h_{2\mathcal{R}}^2)$. We define another variable $\bar{h} = \min(h_{1\mathcal{R}}^2, h_{2\mathcal{R}}^2)$ and we have $m_{L_{\oplus}} = 2\bar{h}\rho$. Also, the PDF of \bar{h} is written as $p_{\bar{h}}(\bar{h}) = \bar{\lambda} \exp(-\bar{\lambda}\bar{h})$, where based on the proof in Lemma 1, we have $\bar{\lambda} = \frac{\lambda_{1\mathcal{R}} + \lambda_{2\mathcal{R}}}{\lambda_{1\mathcal{R}} \lambda_{2\mathcal{R}}}$.

We define the diversity as

$$D \triangleq - \lim_{\rho \rightarrow \infty} \frac{\log \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}, h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{R}\mathcal{D}}} (P_{e,i})}{\log \rho}, \quad (57)$$

where

$$\begin{aligned} \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}, h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{R}\mathcal{D}}} (P_{e,i}) &= \mathbb{E}(P_{i,mrc}) \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_c)) \\ &+ \mathbb{E}(P_{i,neg}) \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e)) + \mathbb{E}(P_{i,dir}) \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_s)). \end{aligned} \quad (58)$$

As we have derived $\mathbb{E}(P_{i,mrc})$, $\mathbb{E}(P_{i,neg})$, and $\mathbb{E}(P_{i,dir})$, we will focus on $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_c))$, $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e))$, and $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_s))$ in the following.

As the optimal threshold can be approximated as $L_T^* \approx \ln \rho$ in the high SNR region, after some manipulations, we calculate the expectation of $\Pr(\mathcal{E}_e)$ in (9) as

$$\begin{aligned} \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e)) &= \int_0^\infty Q \left(\frac{\ln \rho}{\sqrt{4h\rho}} + \sqrt{h\rho} \right) p_h(h) dh \\ &= \underbrace{\int_0^{\frac{2\ln \rho}{\rho}} Q \left(\frac{\ln \rho}{\sqrt{4h\rho}} + \sqrt{h\rho} \right) p_h(h) dh}_{\mathcal{A}} \\ &+ \underbrace{\int_{\frac{2\ln \rho}{\rho}}^\infty Q \left(\frac{\ln \rho}{\sqrt{4h\rho}} + \sqrt{h\rho} \right) p_h(h) dh}_{\mathcal{B}}. \end{aligned} \quad (59)$$

We first note that $\frac{\ln \rho}{\sqrt{4h\rho}} + \sqrt{h\rho} \geq 2\sqrt{\frac{\ln \rho}{4h\rho} \sqrt{h\rho}} = 2\sqrt{\frac{1}{2} \ln \rho}$. Then for the function \mathcal{A} in (59), when ρ is large, we have

$$\begin{aligned} \mathcal{A} &\leq Q \left(2\sqrt{\frac{1}{2} \ln \rho} \right) \int_0^{\frac{2\ln \rho}{\rho}} p_h(h) dh \\ &\stackrel{(a)}{\leq} \frac{1}{2} \exp(-\ln \rho) \left(1 - \exp\left(-\frac{2\ln \rho}{\rho}\right) \right) \\ &\stackrel{(b)}{\approx} \frac{\ln \rho}{\rho^2}, \end{aligned} \quad (60)$$

where (a) in (60) follows the Chernoff bound of $Q(\cdot)$, i.e., $Q(z) \leq \frac{1}{2} \exp\left(-\frac{z^2}{2}\right)$, and (b) applies the Taylor expansion $\exp(z) \approx 1 + z$ when z approaches zero.

For the function \mathcal{B} in (59), we have

$$\begin{aligned} \mathcal{B} &\leq \int_{\frac{2\ln \rho}{\rho}}^\infty Q \left(\sqrt{h\rho} \right) p_h(h) dh \\ &\leq \frac{\bar{\lambda}}{2} \int_{\frac{2\ln \rho}{\rho}}^\infty \exp\left(-\frac{h\rho}{2} - \bar{\lambda}h\right) dh \\ &= \frac{\bar{\lambda}}{2} \frac{1}{\frac{\rho}{2} + \bar{\lambda}} \exp\left(-\ln \rho - \frac{2\ln \rho}{\rho} \bar{\lambda}\right). \end{aligned} \quad (61)$$

When ρ goes to infinity, by using base-2 logarithm, we obtain

$$\begin{aligned} - \lim_{\rho \rightarrow \infty} \frac{\log \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e))}{\log \rho} &= - \lim_{\rho \rightarrow \infty} \frac{\log(\mathcal{A} + \mathcal{B})}{\log \rho} \\ &= 2. \end{aligned} \quad (62)$$

Next, we investigate the expectation of $\Pr(\mathcal{E}_s)$ in (9), which can be rewritten as

$$\begin{aligned} \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_s)) &= \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} \left(\int_{-\infty}^{L_T} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL \right) - \\ &\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} \left(\int_{-\infty}^{-L_T} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL \right). \end{aligned} \quad (63)$$

Note that the second item in the right hand side of (63) is $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e))$. Therefore, we only focus on the derivation of the first item, which can be calculated as

$$\begin{aligned} \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} \left(\int_{-\infty}^{L_T} p_{L_{x_{\mathcal{R}}^j, \mathcal{R}}}(L) dL \right) &= \\ &\underbrace{\int_0^{\frac{\ln \rho}{2\rho}} \left(1 - Q \left(\frac{\ln \rho}{\sqrt{4h\rho}} - \sqrt{h\rho} \right) \right) p_h(h) dh}_{\mathcal{C}} + \\ &\underbrace{\int_{\frac{\ln \rho}{2\rho}}^{\frac{\ln \rho}{\rho}} Q \left(\sqrt{h\rho} - \frac{\ln \rho}{\sqrt{4h\rho}} \right) p_h(h) dh}_{\mathcal{D}} + \\ &\underbrace{\int_{\frac{\ln \rho}{\rho}}^\infty Q \left(\sqrt{h\rho} - \frac{\ln \rho}{\sqrt{4h\rho}} \right) p_h(h) dh}_{\mathcal{F}}. \end{aligned} \quad (64)$$

It is easy to obtain that

$$\begin{aligned} \mathcal{C} + \mathcal{D} &\leq \int_0^{\frac{\ln \rho}{\rho}} p_h(h) dh \\ &= 1 - \exp\left(-\frac{\ln \rho}{\rho}\right) \\ &\approx \frac{\ln \rho}{\rho}. \end{aligned} \quad (65)$$

Regarding the third item in the right hand side of (64), i.e., the function \mathcal{F} , we have

$$\begin{aligned} \mathcal{F} &\leq \int_{\frac{\ln \rho}{\rho}}^\infty Q \left(\frac{1}{2} \sqrt{h\rho} \right) p_h(h) dh \\ &\leq \frac{\bar{\lambda}}{2} \int_{\frac{\ln \rho}{\rho}}^\infty \exp\left(-\frac{h\rho}{8} - \bar{\lambda}h\right) dh \\ &= \frac{\bar{\lambda}}{2} \frac{1}{\frac{\rho}{8} + \bar{\lambda}} \exp\left(-\frac{\ln \rho}{8} - \frac{\ln \rho}{2\rho} \bar{\lambda}\right). \end{aligned} \quad (66)$$

When ρ goes to infinity, we obtain

$$\begin{aligned} - \lim_{\rho \rightarrow \infty} \frac{\log \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_s))}{\log \rho} &= - \lim_{\rho \rightarrow \infty} \frac{\log(\mathcal{C} + \mathcal{D} + \mathcal{F})}{\log \rho} \\ &= 1. \end{aligned} \quad (67)$$

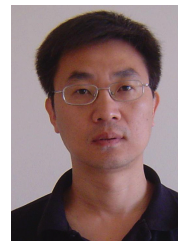
Combining $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_e))$ and $\mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}} (\Pr(\mathcal{E}_s))$ with the equations in (27) and (28), we thus obtain that $-\lim_{\rho \rightarrow \infty} \frac{\log \mathbb{E}_{h_{1\mathcal{R}}, h_{2\mathcal{R}}, h_{1\mathcal{D}}, h_{2\mathcal{D}}, h_{\mathcal{R}\mathcal{D}}} (P_{e,i})}{\log \rho} = 2$. Similarly, we can obtain the same result for $x_1^j x_2^j = -1$, and thus each source can achieve a diversity of two. This completes the proof. ■

REFERENCES

- [1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [3] D. Chen, K. Azarian, and J. N. Laneman, "A case for amplify-forward relaying in the block-fading multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3728–3733, Aug. 2008.
- [4] T. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 375–391, Feb. 2006.
- [5] T. Hunter and A. Nosratinia, "Diversity through coded cooperation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 283–289, Feb. 2006.
- [6] Y. Zhang and Z. Zhang, "Joint network-channel coding with rateless code over multiple access relay system," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 320–332, Jan. 2013.
- [7] J. Li, J. Yuan, R. Malaney, M. H. Azmi, and M. Xiao, "Network coded LDPC code design for a multi-source relaying system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1538–1551, May 2011.
- [8] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Proc. 2006 IEEE Conference on Sensor and Ad Hoc Communications and Networks*, pp. 817–822.
- [9] J. Li, J. Yuan, R. Malaney, M. Xiao, and W. Chen, "Full-diversity binary frame-wise network coding for multiple-source multiple-relay networks over slow-fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1346–1360, Mar. 2012.
- [10] M. Xiao, J. Kliewer, and M. Skoglund, "Design of network codes for multiple-user multiple-relay wireless networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3755–3766, Dec. 2012.
- [11] M. Xiao and M. Skoglund, "Multiple-user cooperative communications based on linear network coding," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3345–3351, Dec. 2010.
- [12] J. Rebelatto, B. Uchoa-Filho, Y. Li, and B. Vucetic, "Multiuser cooperative diversity through network coding based on classical coding theory," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 916–926, Feb. 2012.
- [13] H. Topakkaya and Z. Wang, "Wireless network code design and performance analysis using diversity-multiplexing tradeoff," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 488–496, Feb. 2011.
- [14] M. Di Renzo, M. Iezzi, and F. Graziosi, "On diversity order and coding gain of multisource multirelay cooperative wireless networks with binary network coding," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1138–1157, Mar. 2013.
- [15] —, "Error performance and diversity analysis of multi-source multi-relay wireless networks with binary network coding and cooperative MRC," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2883–2903, Jun. 2013.
- [16] A. Nasri, R. Schober, and M. Uysal, "Performance and optimization of network-coded cooperative diversity systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1111–1122, Mar. 2013.
- [17] Y. Li, B. Vucetic, T. Wong, and M. Dohler, "Distributed turbo coding with soft information relaying in multihop relay networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp. 2040–2050, Nov. 2006.
- [18] K. S. Gomadam and S. A. Jafar, "Optimal relay functionality for SNR maximization in memoryless relay networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 390–401, Feb. 2007.
- [19] J. Li, M. Karim, J. Yuan, Z. Chen, Z. Lin, and B. Vucetic, "Novel soft information forwarding protocols in two-way relay channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2374–2381, Jun. 2013.
- [20] M. A. Karim, J. Yuan, Z. Chen, and J. Li, "Soft information relaying in fading channels," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 233–236, Jun. 2012.
- [21] S. Tian, Y. Li, and B. Vucetic, "Piecewise-and-forward relaying in wireless relay networks," *IEEE Signal Process. Lett.*, vol. 18, no. 5, pp. 323–326, May 2011.
- [22] H. V. Khuong and H. Y. Kong, "LLR-based decode-and-forward protocol for relay networks and closed-form BER expressions," *IEICE Trans. Fundamentals*, vol. E89A, no. 6, pp. 1832–1841, Jun. 2006.
- [23] G. Al-Habian, A. Ghayeb, M. Hasna, and A. Abu-Dayya, "Threshold-based relaying in coded cooperative networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 123–135, 2011.
- [24] S. L. H. Nguyen, A. Ghayeb, G. Al-Habian, and M. Hasna, "Mitigating error propagation in two-way relay channels with network coding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3380–3390, Nov. 2010.
- [25] S. Yang and R. Koetter, "Network coding over a noisy relay: a belief propagation approach," in *Proc. 2007 IEEE International Symposium on Information Theory*, pp. 801–804.
- [26] G. Zeitler, G. Bauch, and J. Widmer, "Quantize-and-forward schemes for the orthogonal multiple-access relay channel," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1148–1158, Apr. 2012.
- [27] W. Guan and K. J. R. Liu, "Mitigating error propagation for wireless network coding," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3632–3643, Oct. 2012.
- [28] T. Wang, G. B. Giannakis, and R. Wang, "Smart regenerative relays for link-adaptive cooperative communications," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1950–1960, Nov. 2008.
- [29] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
- [30] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.
- [31] L. Lu and S.-C. Liew, "Asynchronous physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 819–831, Feb. 2012.
- [32] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657–670, Feb. 2001.
- [33] S. Ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
- [34] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, Sep. 1998.
- [35] F. Lehmann and G. M. Maggio, "Analysis of the iterative decoding of LDPC and product codes using the Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2993–3000, Nov. 2003.
- [36] S. Zhang, Y. Zhu, and S. C. Liew, "Soft network coding in wireless two-way relay channels," *J. Commun. and Networks*, vol. 10, no. 4, pp. 371–383, Dec. 2008.
- [37] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Process*, 4th ed. McGraw Hill, 2002.



Jun Li (M'09) received Ph. D degree in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, P. R. China in 2009. From January 2009 to June 2009, he worked in the Department of Research and Innovation, Alcatel Lucent Shanghai Bell as a Research Scientist. From June 2009 to April 2012, he was a Postdoctoral Fellow at the School of Electrical Engineering and Telecommunications, the University of New South Wales, Australia. From April 2012 to now, he is a Research Fellow at the School of Electrical Engineering, the University of Sydney, Australia. He served as the Technical Program Committee member for several international conferences such as APCC2009, APCC2010, VTC2011 (Spring), ICC2011, TENCON2012, APCC2013, VTC2014 (Fall), and ICC2014. His research interests include network information theory, channel coding theory, wireless network coding and cooperative communications.

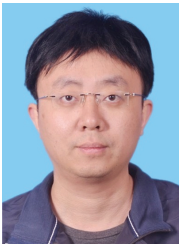


Zihuai Lin received the Ph.D. degree in Electrical Engineering from Chalmers University of Technology, Sweden, in 2006. Prior to this he has held positions at Ericsson Research, Stockholm, Sweden. Following Ph.D. graduation, he worked as a Research Associate Professor at Aalborg University, Denmark and currently at the School of Electrical and Information Engineering, the University of Sydney, Australia. His research interests include graph theory, source/channel/network coding, coded modulation, MIMO, OFDMA, SC-FDMA, radio resource management, cooperative communications, small-cell networks, 5G cellular systems, etc.

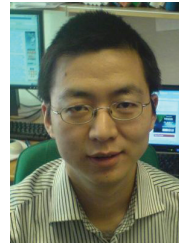


Branka Vucetic (F'03) received the B.S.E.E., M.S.E.E., and Ph.D. degrees in electrical engineering, from the University of Belgrade, Belgrade, Yugoslavia, in 1972, 1978, and 1982, respectively. She currently holds the Peter Nicol Russel Chair of Telecommunications Engineering at the University of Sydney. She is an internationally recognized expert in wireless communications and coding and has been elected to the IEEE Fellow grade for her research contributions in channel coding and its applications in wireless communications. Her research

has involved collaborations with industry and government organizations in Australia and several other countries.



Jin Xu received the Ph.D. degree in electrical engineering from the Wuhan University, P. R. China in 2007. He is currently with ZTE Corp., where he is a senior engineer in the wireless pre-research team. His research interests include channel coding, network coding, multiple-access communications as well as their application in cellular mobile networks.



Ming Xiao (S'02-M'07-SM'12) was born in the SiChuan Province, P. R. China, in 1975. He received Bachelor and Master degrees in Engineering from the University of Electronic Science and Technology of China, ChengDu in 1997 and 2002, respectively. He received Ph.D degree from Chalmers University of technology, Sweden in November 2007. From 1997 to 1999, he worked as a network and software engineer in ChinaTelecom. From 2000 to 2002, he also held a position in the SiChuan communications administration. From November 2007 to now, he

has been in Communication Theory, school of electrical engineering, Royal Institute of Technology, Sweden, where he is currently an Associate Professor (and Docent) in Communications Theory. He received the best paper Awards in IC-WCSP (International Conference on Wireless Communications and Signal Processing) in 2010 and IEEE ICCCN (International Conference on Computer Communication Networks) in 2011. Dr. Xiao received "Chinese Government Award for Outstanding Self-Financed Students Studying Abroad" in March, 2007. He received a "Hans Werthen Grant" from the Royal Swedish Academy of Engineering Science (IVA) in March 2006. He received "Ericsson Research Funding" from Ericsson in 2010. Since 2012, he has been an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, and IEEE WIRELESS COMMUNICATIONS LETTERS.



Wen Chen (M'03-SM'11) received the B.S. and M.S. degrees from Wuhan University, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999. He was a researcher of Japan Society for the Promotion of Sciences (JSPS) from 1999 through 2001. In 2001, he joined the University of Alberta, Canada, starting as a Postdoctoral Fellow with the Information Research Laboratory and continuing as a Research Associate in the Department of Electrical and Computer Engineering. Since 2006,

he has been a Full Professor in the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Institute for Signal Processing and Systems. His interests cover network coding, cooperative communications, cognitive radio, and MIMO-OFDM systems.