A Lower Bound Analysis of Hamming Distortion for a Binary CEO Problem with Joint Source-Channel Coding

Xin He, Xiaobo Zhou, *Member, IEEE,* Petri Komulainen, *Member, IEEE,* Markku Juntti, *Senior Member, IEEE* and Tad Matsumoto *Fellow, IEEE*

***Abstract*—A two-node binary chief executive officer (CEO) problem is investigated. Noise-corrupted versions of a binary sequence are forwarded by two nodes to a single destination node over orthogonal additive white Gaussian noise (AWGN) channels. We first reduce the binary CEO problem to a binary multiterminal source coding problem, of which an outer bound for the rate-distortion region is derived. The distortion function is then established by evaluating the relationship between the binary CEO and multiterminal source coding problems. A lower bound approximation on the Hamming distortion (HD) is obtained by minimizing a distortion function subject to constraints obtained based on the source-channel separation theorem. Encoding/decoding algorithms using concatenated convolutional codes and a joint decoding scheme are used to verify the lower bound on the HD. It is found that the theoretical lower bounds on the HD and the computer simulation based bit error rate performance curves have the same tendencies. The differences in the threshold signal-to-noise ratio between the theoretical lower bounds and those obtained by simulations are around 1.5 dB in AWGN channel. The theoretical lower bound on the HD in block Rayleigh fading channel is also evaluated by performing Monte Carlo simulation.**

***Index Terms*—Hamming distortion lower bound, binary CEO problem, binary multiterminal source coding, rate-distortion outer bound.**

I. Introduction

THE chief executive officer (CEO) problem where a CEO aims at reproducing a common source which cannot be directly observed was first introduced by Berger *et al.* in [1]. The CEO problem has attracted a lot of attention, not only from pure information theoretic interest but also from practical application viewpoint such as wireless sensor networks (WSNs).

The quadratic Gaussian CEO problem, a particular case of the CEO problem, where the source and the multiple observations are assumed to be jointly Gaussian distributed, was studied in [2], [3], where an explicit form of the rate-distortion function was derived. Chen *et al.* derived an upper bound on the sum-rate for the CEO problem and proposed rate allocation schemes by exploiting the contra-polymatroid structure of the achievable rate region in [4]. Besides the theoretical work, the performance on minimum achievable distortion of a successive coding strategy [5] based on a generalization of Wyner-Ziv source coding was evaluated in [6], and the optimal rate allocation scheme to achieve the minimum distortion under a sum-rate constraint was further proposed in [7] for the successively structured Gaussian CEO problem.

In this work, we focus on a binary CEO problem, in which a binary source is estimated by the CEO through multiple deployed nodes. There are many practical applications of the binary CEO problem, for example, a binary data gathering sensor network, distributed detection using multiple sensors [8], power-distortion tradeoff in sensor networks [9], and wireless mesh network (WMN) with *lossy forwarding* (LF) [10].

An iterative joint decoding algorithm for the binary data gathering WSN, which is a direct application of the binary CEO problem, was proposed in [11], where a convolutional code is applied at the sensor node. A coding scheme based on the parallel concatenated convolutional codes was proposed in [12], where the extrinsic log-likelihood ratios (LLRs) are weighted by the observation error probabilities at the decoder. Moreover, the capacity of the equivalent parallel channel was derived to verify the bit error rate (BER) performance taking into account the error probability of the observed data sequence. In [13], an adaptive bi-modal decoder for a binary source estimation involving two sensors was proposed based on the modified extrinsic information transfer (EXIT) chart analysis. A convergence property analysis of the iterative decoding algorithm was presented based on the modified EXIT chart analysis for binary data gathering WSNs in [14]. It shows that iterative process is less useful if the channel quality is very good or the observation accuracy is very low. In [10], we proposed an encoding/decoding technique which can significantly improve the BER performance by exploiting the correlation

This work has been in part supported by Academy of Finland, and in part performed in the framework of the FP7 project ICT-619555 RESCUE (Links-on-the-fly Technology for Robust, Efficient and Smart Communication in Unpredictable Environments) which is partly funded by the European Union.

X. He and T. Matsumoto are with the School of Information Science, JAIST, 1-1 Asahidai, Nomi, Ishikawa, Japan 923-1292 (e-mail: {hexin,matumoto}@jaist.ac.jp) and the Centre for Wireless Communications and Department of Communications Engineering, University of Oulu, FI-90014 Finland (e-mail: {xin.he, tadashi.matsumoto}@ee.oulu.fi).

X. Zhou was with the Centre for Wireless Communications and Department of Communications Engineering, University of Oulu and is now with School of Computer Science and Technology, Tianjin University, Tianjin, China (email:xiaobo.zhou@tju.edu.cn).

P. Komulainen was with the Centre for Wireless Communications and Department of Communications Engineering, University of Oulu and is now with MediaTek, Oulu, Finland (e-mail: petri.komulainen@mediatek.com).

M. Juntti is with the Centre for Wireless Communications and Department of Communications Engineering, University of Oulu, FI-90014 Finland (e-mail: markku.juntti@oulu.fi).

knowledge through the LLR updating function [15], for both a binary independently and identically distributed (i.i.d.) source and a binary Markov source. In [16], we proposed a non-negative constrained iterative algorithm for estimating the observation error probabilities in a WSN having an arbitrary number of sensors.

For the binary CEO problem, most of the previous works focus on the design of practical encoding/decoding algorithms. This motivates us to analyze a problem that how small a distortion level the CEO can achieve from the rate-distortion perspective. Hence, the primary goals of this work are to theoretically provide a lower bound on the Hamming distortion (HD), corresponding to the bit error probability (BEP), and to verify the lower bound by making performance comparison with a practical encoding/decoding algorithm.

The major contribution is the derivation of a HD approximation based on an information-theoretic outer bound resulting in a lower bound (approximation) on the HD (or BEP) for a two-node communication network with joint source-channel (JSC) coding. Its objective is to estimate a single binary source over two orthogonal additive white Gaussian noise (AWGN) channels. For the simplicity of analysis, we focus on orthogonal transmissions from two nodes to the destination, and we sperate the stages for JSC decoding and the final decision on the common source [14], [17], [18]. Hence, deriving the theoretical lower bound on the HD is equivalent to minimizing a distortion function subject to a series of inequalities obtained based on the source-channel separation theorem for lossy source coding.

In order to solve the minimization problem, we first model the source coding of this two-node network by the binary CEO problem. We then reduce the binary CEO problem to a binary multiterminal source coding problem, which plays the core role in solving the main problem. An outer bound for the rate-distortion region of the binary multiterminal source coding problem is then derived by providing the converse proof. We establish the relationship between the binary CEO problem and the binary multiterminal source coding problem in terms of the distortion function. Finally, the minimization problem is formulated in the framework of convex optimization. It should be emphasized here that our purpose is not to derive a tight rate-distortion bound for the binary CEO problem. Instead, we focus on the derivation of a lower bound that can be used as a reference of the BER performance curves of the encoding/decoding algorithms, including the technique proposed in [10] and [16].

The rest of the paper is organized as follows. In Section II, the system model and the problem to be solved are described. The derivation of the outer bound and its proof for the binary multiterminal source coding problem is detailed in Section III. The problem of how to obtain the lower bound on the HD is formulated in Section IV. Section V provides the numerical results of simulations as well as their corresponding lower bounds. Finally, we conclude this work in Section VI with several concluding statements.

II. Problem Statement

Notation. The uppercase and lowercase letters are used to denote random variables and their realizations, respectively. The alphabet set of a random variable X is denoted by $\mathcal{X}$. Let $\mathbf{X}^n$ and $\mathbf{x}^n$ represent a random vector and its realization, respectively, with the superscript n being the length of the vector (block length). We use t to denote the time index and i to denote the index of a node.

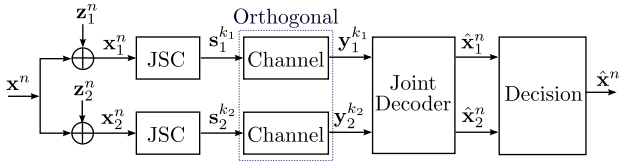

Fig. 1. The abstract system model of estimating a single source through two independent nodes with joint source-channel coding.

The system model of estimating a single source through two nodes is depicted in Fig. 1. A common i.i.d. source X produces a sequence $\mathbf{x}^n = \{x(t)\}_{t=1}^n$ by taking values from a binary set $\mathcal{X} = \{0, 1\}$ with equal probability. Source X is observed by two nodes and forwarded to a single destination. Due to the inaccuracy of the estimation and/or limited received signal power at nodes, such as in WSN and WMN, the sequences received by the nodes may contain errors[1], and the nodes still forward the erroneous sequences to the destination, which is referred as LF [20], [21]. The error probabilities $\Pr(x_1(t) \neq x(t))$ and $\Pr(x_2(t) \neq x(t))$ are denoted as p_1 and p_2, respectively, i.e., $\Pr(z_i(t) = 1) = p_i$ for the binary noise sequence $\mathbf{z}_i^n = \{z_i(t)\}_{t=1}^n$, $i = 1, 2$. At the nodes, the noisy versions $\mathbf{x}_1^n = \{x_1(t)\}_{t=1}^n$ and $\mathbf{x}_2^n = \{x_2(t)\}_{t=1}^n$ of $\mathbf{x}^n$ are separately encoded by two JSC encoders to generate symbol sequences $\mathbf{s}_1^{k_1} = \{s_1(t)\}_{t=1}^{k_1}$ and $\mathbf{s}_2^{k_2} = \{s_2(t)\}_{t=1}^{k_2}$ with coding rates $r_i = n/k_i$, $i = 1, 2$. The symbol sequences $\mathbf{s}_1^{k_1}$ and $\mathbf{s}_2^{k_2}$ are then transmitted to the destination over two orthogonal AWGN channels, as

$$\mathbf{y}_i^{k_i} = h_i \cdot \mathbf{s}_i^{k_i} + \mathbf{w}_i^{k_i}, i = 1, 2, \tag{1}$$

where h_i and $\mathbf{w}_i^{k_i} = \{w(t)\}_{t=1}^{k_i}$ represent the channel gain and the AWGN sequence at the destination, respectively. The orthogonality can be achieved by any scheduled multiple access scheme, like time division multiple access (TDMA), i.e., $\mathbf{s}_1^{k_1}$ and $\mathbf{s}_2^{k_2}$ can be transmitted at different time intervals. The destination performs JSC decoding to form estimates $\hat{\mathbf{x}}_i^n$ of the sequences $\mathbf{x}_i^n$, $i = 1, 2$. We define the expected Hamming distortion measures $E[\frac{1}{n}\sum_{t=1}^n d(x_i(t), \hat{x}_i(t))] \leq D_i + \epsilon$ to evaluate the error probability $\Pr(x_i(t) \neq \hat{x}_i(t))$ with

$$d(x_i(t), \hat{x}_i(t)) = \begin{cases} 1, & \text{if } x_i(t) \neq \hat{x}_i(t), \\ 0, & \text{if } x_i(t) = \hat{x}_i(t), \end{cases} \tag{2}$$

and ϵ representing an arbitrarily small positive number.

Finally, the destination reconstructs the source information $\mathbf{x}^n$ of which the estimate is denoted as $\hat{\mathbf{x}}^n$ based on a decision rule from $\hat{\mathbf{x}}_1^n$ and $\hat{\mathbf{x}}_2^n$. Therefore, the distortion measure $E[\frac{1}{n}\sum_{t=1}^n d(x(t), \hat{x}(t))] \leq D + \epsilon$ can be formulated as a

[1]In WMN applications, the nodes correspond to the transceivers in the multiple routes. In a WMN, a source communicates with a destination through multiple intermediate nodes if they are not within the communication coverage. If errors are allowed in the messages forwarded by the intermediate nodes, the WMN can be also modeled as the model shown in Fig. 1 [19].

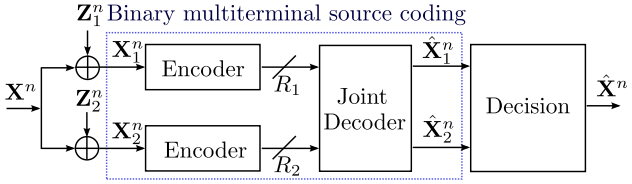

Fig. 2. The abstract model of the binary CEO problem with two independent nodes.

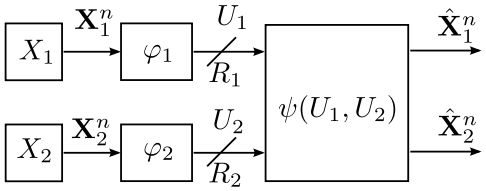

Fig. 3. The binary multiterminal source coding problem for two correlated binary sources.

function of D_i, $i = 1, 2$, as $D = f(D_1, D_2)$, where function $f(\cdot)$ is detailed in subsection IV-A. It should be emphasized here that function $D = f(D_1, D_2)$ limits the decoding scheme to which first reconstructs $\mathbf{x}_1^n$ and $\mathbf{x}_2^n$ and then makes the decision on $\mathbf{x}^n$ from those reconstructions (it is referred to as sequential decoding), as shown in Fig. 1. The optimality of such a decoding scheme is an open problem, but it is definitely of interest for practical systems. Furthermore, $f(D_1, D_2)$ largely depends on the decision rule, i.e., there exists different function $f(D_1, D_2)$ for different decision rules[2].

According to the source-channel separation theorem for lossy source coding [23], distortions D_1 and D_2 can be achieved if the following inequalities hold:

$$\begin{cases} R_1(D_1) \cdot r_1 & \leq C(\gamma_1), \\ R_2(D_2) \cdot r_2 & \leq C(\gamma_2), \end{cases} \tag{3}$$

where $R_i(D_i)$ is the rate-distortion function for the source coding and $C(\gamma)$ is the Shannon capacity using Gaussian codebook[3] with the argument γ denoting the signal-to-noise ratio (SNR) of the channel. As stated above, r_1 and r_2 represent end-to-end coding rates of two links. Our goal is to derive the theoretical lower bound on the HD for the system shown in Fig. 1. It is equivalent to minimizing the expected Hamming distortion D through a function $f(D_1, D_2)$ under constraints shown in (3), as

$$\begin{aligned} \min_{D_1, D_2} \quad & D = f(D_1, D_2) \\ \text{s.t.} \quad & (3). \end{aligned} \tag{4}$$

The minimization being performed in (4) is for a specific system which maps the average distortions D_1 and D_2 to D, since function $f(D_1, D_2)$ is defined for designated decision rules. To achieve this goal by solving (4), we turn to derive the rate-distortion function $R_i(D_i)$ for the problem shown in Fig. 1 and to establish the function $D = f(D_1, D_2)$ for the decision rule used at the destination.

III. Rate-Distortion Region Analysis

A. Source Coding

In network information theory, the source coding of the communication system shown in Fig. 1 is modeled by the binary CEO problem. The abstract model of the binary CEO problem is illustrated in Fig. 2. In order to derive the rate-distortion function $R_i(D_i)$, we first reduce the binary CEO problem to a binary multiterminal source coding problem. An outer bound for the rate-distortion region which is determined by the rate-distortion function $R_i(D_i)$ is then derived for the binary multiterminal source coding problem through the converse proof, as in the Gaussian case [25].

The binary multiterminal source coding problem which we consider is depicted in Fig. 3. Since random sources $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$ originate from the common source $\mathbf{X}^n$, the random variable pair (X_1, X_2) follows a joint probability distribution $\Pr_{X_1,X_2}(x_1, x_2)$ given by

$$\Pr_{X_1,X_2}(x_1, x_2) = \begin{cases} \frac{1}{2} \cdot \rho, & \text{if } x_1 \neq x_2, \\ \frac{1}{2} \cdot (1 - \rho), & \text{otherwise}, \end{cases} \tag{5}$$

where $\rho = \Pr(x_1 \neq x_2)$ is the correlation parameter between the sources X_1 and X_2, i.e., X_2 can be seen as the output of a binary symmetric channel (BSC) with the crossover probability ρ where X_1 is the input. Two encoders separately encode $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$ at rates R_1 and R_2 as

$$\begin{aligned} \varphi_1 : & \mathcal{X}^n \to \mathcal{M}_1 = \{1, 2, \cdots, 2^{nR_1}\}, \\ \varphi_2 : & \mathcal{X}^n \to \mathcal{M}_2 = \{1, 2, \cdots, 2^{nR_2}\}. \end{aligned}$$

The encoder output sequences $U_1 = \varphi_1(\mathbf{X}_1^n)$ and $U_2 = \varphi_2(\mathbf{X}_2^n)$ are transmitted to a common receiver. It jointly decodes the received samples to construct the estimates $(\hat{\mathbf{X}}_1^n, \hat{\mathbf{X}}_2^n)$ of the source pair $(\mathbf{X}_1^n, \mathbf{X}_2^n)$ denoted as $(\hat{\mathbf{X}}_1^n, \hat{\mathbf{X}}_2^n) = \psi(\varphi_1(\mathbf{X}_1^n), \varphi_2(\mathbf{X}_2^n))$.

For given distortion values $D_1 \in [0, \frac{1}{2}]$ and $D_2 \in [0, \frac{1}{2}]$, the rate-distortion region $\mathcal{R}(D_1, D_2)$ is defined as

$$\begin{aligned} \mathcal{R}(D_1, D_2) = \Big\{ & (R_1, R_2) : (R_1, R_2) \text{ is admissible such that} \\ & E\frac{1}{n}\sum_{t=1}^{n} d(x_i(t), \hat{x}_i(t)) \leq D_i + \epsilon, i = 1, 2 \Big\}. \end{aligned}$$

B. Outer Bound for Rate-Distortion Region

We provide a bound $\mathcal{R}^{\text{o}}(D_1, D_2)$ of the rate-distortion region $\mathcal{R}(D_1, D_2)$.

Definition 1:

$$\begin{aligned} \mathcal{R}_1^{\text{o}}(D_1) = \Big\{ & (R_1, R_2) : \forall R_2 \leq 1 \\ & R_1 \geq H_b[\rho * H_b^{-1}(1 - R_2)] - H_b(D_1) \Big\}, \end{aligned} \tag{6}$$

[2]It has been assumed in this setup that 1) each encoder uses joint typicality encoding and binning based on random coding arguments, and the decoder performs joint typicality decoding with a sufficiently large n to achieve the average distortion D_i as in the Berger-Tung source coding problem [22]; 2) the errors occurring in each sequence $\mathbf{x}_i^n$ are i.i.d. In the practical system, we use random interleavers to asymptotically make this assumption practical. As shown in section V-B, the simulation results are consistent with the lower bound calculation based on majority logic decision.

[3]For one dimensional signal, $C(\gamma) = \frac{1}{2}\log_2(1 + 2\gamma)$, and for two dimensional signal, $C(\gamma) = \log_2(1 + \gamma)$ [24].

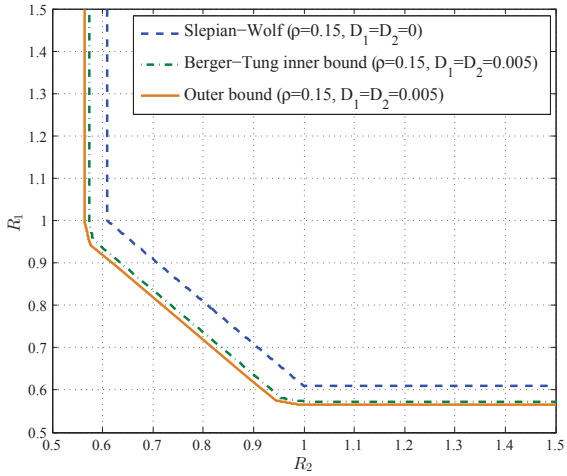

Fig. 4. The comparison of $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$, Berger-Tung inner bound and Slepian-Wolf admissible rate region. The correlation ρ between two sources is set at 0.15.

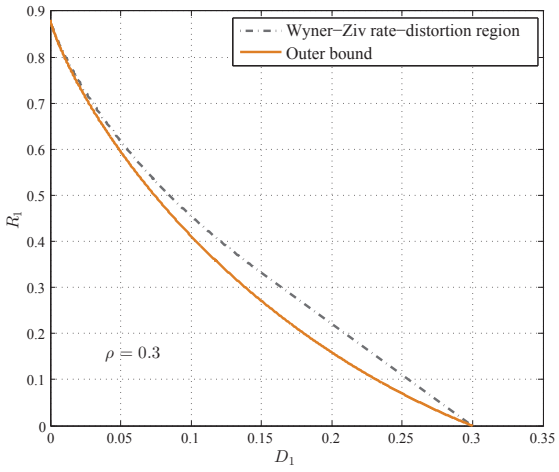

Fig. 5. The comparison of Wyner-Ziv rate-distortion region and derived outer bound. The correlation ρ between two sources is set at 0.3.

with $H_b(a) = -a\log_2(a) - (1-a)\log_2(1-a)$ and $H_b^{-1}(a)$ representing the binary entropy function and its inverse function, respectively[4]. The operator $*$ calculates the binary convolution of the two variables, i.e., $a * b = a(1-b) + b(1-a)$.

$$\mathcal{R}_2^{\mathrm{o}}(D_2) = \Big\{(R_1, R_2) : \forall R_1 \le 1 \\ R_2 \ge H_b[\rho * H_b^{-1}(1-R_1)] - H_b(D_2)\Big\}, \quad (7)$$

$$\mathcal{R}_{12}^{\mathrm{o}}(D_1,D_2) = \Big\{(R_1, R_2) : \\ R_1 + R_2 \ge 1 + H_b(\rho) - H_b(D_1) - H_b(D_2)\Big\}. \quad (8)$$

For every $D_1 \in [0, \frac{1}{2}]$ and $D_2 \in [0, \frac{1}{2}]$,

$$\mathcal{R}^{\mathrm{o}}(D_1, D_2) = \mathcal{R}_1^{\mathrm{o}}(D_1) \bigcap \mathcal{R}_2^{\mathrm{o}}(D_2) \bigcap \mathcal{R}_{12}^{\mathrm{o}}(D_1, D_2). \quad (9)$$

In what follows, we prove that $\mathcal{R}_1^{\mathrm{o}}(D_1)$, $\mathcal{R}_2^{\mathrm{o}}(D_2)$ and $\mathcal{R}_{12}^{\mathrm{o}}(D_1, D_2)$ are the supersets of the regions of $\mathcal{R}(D_1, D_2)$. It means that the following theorem holds.

Theorem 1: $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ is an outer bound for the rate-distortion region $\mathcal{R}(D_1, D_2)$; i.e., $\mathcal{R}(D_1, D_2) \subseteq \mathcal{R}^{\mathrm{o}}(D_1, D_2)$.

C. Proof

Proof of Theorem 1 (Converse): To prove Theorem 1, the following three different cases of the binary multiterminal source coding problem are considered.

Case 1. In order to prove that $\mathcal{R}(D_1, D_2) \subseteq \mathcal{R}_1^{\mathrm{o}}(D_1)$, we assume that the rate pair $(R_1, R_2) \in \mathcal{R}(D_1, D_2)$ and show that this implies that $(R_1, R_2) \in \mathcal{R}_1^{\mathrm{o}}(D_1)$. In the proof, $\mathbf{X}_2^n$ is first reconstructed without constraint on D_2 which results in (15), and then $\hat{\mathbf{X}}_2^n$ is regarded as the side information to recover $\mathbf{X}_1^n$. Assume that a rate pair (R_1, R_2) achieves distortion D_1, then

$$\begin{aligned} n \cdot (R_1 + \epsilon) &\ge H(U_1) \\ &\ge H(U_1|U_2) && (10)\\ &= I(\mathbf{X}_1^n; U_1|U_2) && (11)\\ &= I(\mathbf{X}_1^n; U_1, U_2) - I(\mathbf{X}_1^n; U_2) && (12)\\ &\ge I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n) - I(\mathbf{X}_1^n; U_2), && (13) \end{aligned}$$

where the steps are justified, because
(10) conditioning reduces entropy,
(11) U_1 is a function of $\mathbf{X}_1^n$,
(12) the chain rule of mutual information,
(13) $\mathbf{X}_1^n \to (U_1, U_2) \to \hat{\mathbf{X}}_1^n$ forms a Markov chain.

Now our aim is to lower bound $I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n)$ and upper bound $I(\mathbf{X}_1^n; U_2)$. Since $I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n) = H(\mathbf{X}_1^n) - H(\mathbf{X}_1^n|\hat{\mathbf{X}}_1^n) = n - H(\mathbf{X}_1^n|\hat{\mathbf{X}}_1^n)$, to lower bound $I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n)$ is equivalent to upper bound $H(\mathbf{X}_1^n|\hat{\mathbf{X}}_1^n)$. According to the Fano's inequality, we have

$$\begin{aligned} H(\mathbf{X}_1^n|\hat{\mathbf{X}}_1^n) &\le n \cdot H_b(D_1) + n \cdot D_1 \cdot \log(|\mathcal{X}| - 1) \\ &= n \cdot H_b(D_1). && (14) \end{aligned}$$

On the other hand, since $I(\mathbf{X}_1^n; U_2) = H(\mathbf{X}_1^n) - H(\mathbf{X}_1^n|U_2) = n - H(\mathbf{X}_1^n|U_2)$, an upper bound on $I(\mathbf{X}_1^n; U_2)$ corresponds to the lower bound on $H(\mathbf{X}_1^n|U_2)$. Observing that $\mathbf{X}_1^n \to \mathbf{X}_2^n \to U_2$ forms a Markov chain, it can be shown that $H(\mathbf{X}_1^n|U_2) \ge nH_b(\rho * \beta)$ by [26, Corollary 4], where $\beta = \frac{1}{n}H_b^{-1}[H_b(\mathbf{X}_2^n|U_2)]$. A more detailed explanation is given in Appendix A. Since the binary convolution $*$ is monotonically increasing with respect to β if ρ is fixed, we need to find the minimizing value of β to lower bound $H_b(\rho * \beta)$[5]. We also have the rate constraint on R_2 as

$$\begin{aligned} n \cdot (R_2 + \epsilon) &\ge H(U_2) \\ &= I(\mathbf{X}_2^n; U_2). && (15) \end{aligned}$$

[4]The inverse function $H_b^{-1}(a) : [0,1] \to [0, \frac{1}{2}]$ only takes values from the interval $[0, \frac{1}{2}]$ since distortion is assumed to be within this range.

[5]The binary entropy function is a monotonically increasing function in the interval $[0, \frac{1}{2}]$.

Letting $\epsilon \to 0$, we have $n \cdot R_2 \geq I(\mathbf{X}_2^n; U_2) = H(\mathbf{X}_2^n) - H(\mathbf{X}_2^n|U_2) = n - nH_b(\beta)$, and hence $\beta \geq H_b^{-1}(1-R_2)$. Therefore, the lower bound on $H(\mathbf{X}_1^n|U_2)$ is given by

$$H(\mathbf{X}_1^n|U_2) \geq n \cdot H_b[\rho * H_b^{-1}(1-R_2)]. \tag{16}$$

From (13), (14) and (16), we can obtain

$$\begin{aligned} n \cdot (R_1 + \epsilon) & \\ &\geq n - n \cdot H_b(D_1) - n + n \cdot H_b[\rho * H_b^{-1}(1-R_2)] \\ &= n \cdot H_b[\rho * H_b^{-1}(1-R_2)] - n \cdot H_b(D_1). \end{aligned} \tag{17}$$

The rate-distortion region $\mathcal{R}_1^{\mathrm{o}}(D_1)$ shown in (6) is obtained by letting $\epsilon \to 0$ in (17). Thus, the rate pair (R_1, R_2) satisfies condition (6); i.e., $(R_1, R_2) \in \mathcal{R}_1^{\mathrm{o}}(D_1)$.

Case 2. The source $\mathbf{X}_1^n$ acts as a helper to reconstruct $\mathbf{X}_2^n$ under the required distortion level D_2. This is the case symmetric with *Case 1*. The rate-distortion region shown in (7) can be proved in the same way as in *Case 1*.

Case 3. Here, we prove that $(R_1, R_2) \in \mathcal{R}(D_1, D_2)$ implies $(R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{o}}(D_1, D_2)$. To this end, assume that the rate pair (R_1, R_2) achieves distortion D_1 for X_1 and D_2 for X_2. In the following proof, decoder ψ jointly reconstructs the sources $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$ under required distortions D_1 and D_2. The following inequalities are obtained:

$$\begin{aligned} n \cdot (R_1 + R_2 + \epsilon) &\geq H(U_1) + H(U_2) \\ &\geq H(U_1, U_2) \\ &= I(\mathbf{X}_1^n, \mathbf{X}_2^n; U_1, U_2) \qquad (18) \\ &= I(\mathbf{X}_1^n; U_1, U_2) + I(\mathbf{X}_2^n; U_1, U_2|\mathbf{X}_1^n) \\ &= I(\mathbf{X}_1^n; U_1, U_2) + I(\mathbf{X}_2^n; \mathbf{X}_1^n, U_1, U_2) \\ &\quad - n \cdot I(X_1; X_2) \\ &\geq I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n) + I(\mathbf{X}_2^n; \hat{\mathbf{X}}_2^n) \\ &\quad - n \cdot I(X_1; X_2), \qquad (19) \end{aligned}$$

where (18) holds since U_i is a function of $\mathbf{X}_i^n$, $i = 1, 2$. Similarly, by utilizing Fano's inequality to lower bound $I(\mathbf{X}_1^n; \hat{\mathbf{X}}_1^n)$ and $I(\mathbf{X}_2^n; \hat{\mathbf{X}}_2^n)$, we have

$$n \cdot (R_1 + R_2 + \epsilon) \geq n + n \cdot H_b(\rho) - n \cdot H_b(D_1) - n \cdot H_b(D_2). \tag{20}$$

Letting $\epsilon \to 0$ in the above inequality, we conclude that (8) holds. That is, $(R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{o}}(D_1, D_2)$.

Through these three cases, it can be concluded that the admissible rate pair $(R_1+\epsilon, R_2+\epsilon) \in \mathcal{R}^{\mathrm{o}}(D_1, D_2)$. Furthermore, the monotonicity of the outer bound $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ [27] implies that $\mathcal{R}^{\mathrm{o}}(D_1, D_2) \subseteq \mathcal{R}^{\mathrm{o}}(D_1 + \epsilon, D_2 + \epsilon)$. Since (R_1, R_2) is admissible, we conclude that $\mathcal{R}(D_1, D_2) \subseteq \mathcal{R}^{\mathrm{o}}(D_1, D_2)$ by letting $\epsilon \to 0$. ∎

Remark 1: If either $R_1 = 0$ or $R_2 = 0$, $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ is consistent with the rate-distortion function $1 - H_b(D_i)$ for the binary source.

Remark 2: $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ reduces to the Slepian-Wolf rate region [28] for correlated binary sources if we set $D_1 \to 0$ and $D_2 \to 0$. The Slepian-Wolf rate region and $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ are shown in Fig. 4. It can be found that by allowing nonzero distortion values, the sources can be further compressed compared to the Slepian-Wolf lossless case.

Remark 3: If we are interested in reconstructing only one source of the two sources, say X_1, and there is no rate limit on describing X_2^n, i.e., $R_2 \geq \frac{1}{n} H(\mathbf{X}_2^n)$, then it is equivalent to the Wyner-Ziv compression problem [22]. Fig. 5 plots the rate-distortion bound of the Wyner-Ziv source coding [29] and our derived outer bound. In this case, $\mathcal{R}_1^{\mathrm{o}}(D_1)$ is not tight, since it can be found from Fig. 5 that the rate-distortion region of the Wyner-Ziv problem lies inside of $\mathcal{R}_1^{\mathrm{o}}(D_1)$.

Remark 4: As it is known that the exact rate-distortion bound of lossy multiterminal source coding problem lies between the Berger-Tung inner and outer bounds [22]. We also derived the rate-distortion region $\mathcal{R}^{\mathrm{i}}(D_1, D_2)$ based on the Berger-Tung inner bound [30] after several steps of elementary calculation in information theory, as

$$\mathcal{R}^{\mathrm{i}}(D_1, D_2) = \mathcal{R}_1^{\mathrm{i}}(D_1) \bigcap \mathcal{R}_2^{\mathrm{i}}(D_2) \bigcap \mathcal{R}_{12}^{\mathrm{i}}(D_1, D_2) \tag{21}$$

with

$$\begin{cases} \mathcal{R}_1^{\mathrm{i}}(D_1) = \{(R_1, R_2) | R_1 \geq H_b(\rho * D_1 * D_2) - H_b(D_1)\}, \\ \mathcal{R}_2^{\mathrm{i}}(D_2) = \{(R_1, R_2) | R_2 \geq H_b(\rho * D_1 * D_2) - H_b(D_2)\}, \\ \mathcal{R}_{12}^{\mathrm{i}}(D_1, D_2) = \{(R_1, R_2) | \\ \qquad R_1 + R_2 \geq 1 + H_b(\rho * D_1 * D_2) - \sum_{i=1}^{2} H_b(D_i)\}, \end{cases}$$

for every $0 \leq D_1, D_2 \leq \frac{1}{2}$. In Fig. 4, the Berger-Tung inner bound for binary case $\mathcal{R}^{\mathrm{i}}(D_1, D_2)$ is also presented as a reference to verify how close the bounds $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ and $\mathcal{R}^{\mathrm{i}}(D_1, D_2)$ are. It can be seen from the figure that they are very close to each other for small values of D_1 and D_2, i.e., the outer bound can be considered as a useful reference in the evaluation of the BER performance, even though there exists a small gap. However, to resolve this gap, further insightful discussions are still needed [25].

Remark 5: The outer bound $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ is obtained by assuming that the code length n is sufficiently large. However, the penalty on the rate-distortion function is unavoidable if the code length is finite [31], where the rate-distortion function taking into account n is derived for a conventional point-to-point communication system. According to [31], the impact of n is not significant when n is large enough. Therefore, this work focuses on the infinite code length, while the outer bound $\mathcal{R}^{\mathrm{o}}(D_1, D_2)$ of finite code length is left as a future study.

In summary, the rate-distortion function $R_i(D_i)$ is given by

$$\begin{cases} R_1(D_1) \geq H_b[\rho * H_b^{-1}(1 - R_2(D_2))] - H_b(D_1), \\ R_2(D_2) \geq H_b[\rho * H_b^{-1}(1 - R_1(D_1))] - H_b(D_2), \\ \sum_{i=1}^{2} R_i(D_i) \geq 1 + H_b(\rho) - \sum_{i=1}^{2} H_b(D_i). \end{cases} \tag{22}$$

IV. Hamming Distortion Lower Bound

A. Function $D = f(D_1, D_2)$

As stated in Section II, distortion D is a function of distortions D_i, $i = 1, 2$. Function $f(D_1, D_2)$ is obtained by evaluating the relationship between the binary CEO and the binary multiterminal source coding problems in terms of distortions, where the model of the relationship is shown in Fig. 6. The estimate $\hat{X}$ is obtained based on the decision rule from the outputs of two parallel BSCs with crossover

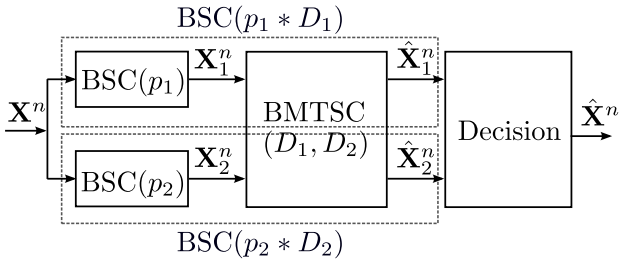

Fig. 6. The distortion model of the binary CEO problem. BMTSC represents binary multiterminal source coding.

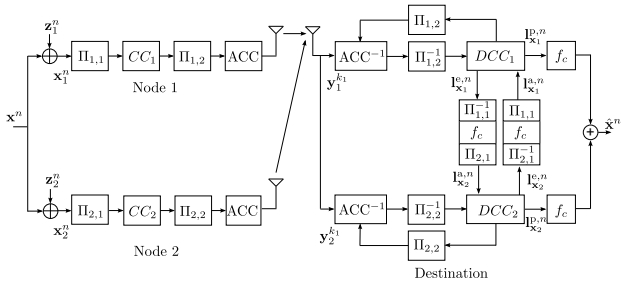

Fig. 7. Block diagram of the encoding/decoding algorithm shown in [10] and [16].

probabilities $p_1 * D_1$, $p_2 * D_2$ and input X. The distortion D largely depends on the decision rule used by the destination. Here we only consider two decision rules. One is the weighted majority decision and the other the optimal decision.

1) Weighted majority decision: Distortion D is obtained by evaluating the probability of an error event. Let $\theta_1 = p_1 * D_1$ and $\theta_2 = p_2 * D_2$. Without loss of generality, we assume that $\theta_1 \leq \theta_2$. Hence, the error event is composed of two independent events: node 1 makes a wrong decision and node 2 makes correct decision or both node 1 and node 2 make erroneous decisions. Therefore, the distortion D in this case is approximated by $D \cong \theta_1(1-\theta_2) + \theta_1\theta_2 = \theta_1$. It can be found that the corner point θ_1 or θ_2 in the rate-distortion region is achieved. Hence, the weighted majority decision rule can be seen as being equivalent to that derived from the time sharing method.

2) Optimal decision: Since the block length is assumed to be infinite and the code is random, an optimal lower bound on the distortion D is determined by utilizing the rate-distortion function for the binary source [24], as

$$\begin{aligned} 1 - H_b(\tilde{d}) &= I(X;\hat{X}) && (23)\\ &\leq I(X;\hat{X}_1,\hat{X}_2)\\ &= H(X) + H(\hat{X}_1,\hat{X}_2) - H(X,\hat{X}_1,\hat{X}_2)\\ &= 1 + 1 + H_b(\theta_1 * \theta_2)\\ &\quad - [H(X) + H(\hat{X}_1|X) + H(\hat{X}_2|X)] && (24)\\ &= 2 + H_b(\theta_1 * \theta_2) - [1 + H_b(\theta_1) + H_b(\theta_2)]\\ &= 1 + H_b(\theta_1 * \theta_2) - H_b(\theta_1) - H_b(\theta_2), && (25) \end{aligned}$$

where $\tilde{d}$ is the Hamming distortion measure between X and $\hat{X}$, and the steps are justified as:
(23) rate-distortion function for the binary source,
(24) $\hat{X}_1 \to X \to \hat{X}_2$ forms a Markov chain.

Thus, it is obvious from (25) that for $0 \leq \tilde{d} \leq \frac{1}{2}$, $\tilde{d} \geq H_b^{-1}[H_b(\theta_1) + H_b(\theta_2) - H_b(\theta_1 * \theta_2)]$. Therefore, the distortion D is the minimum value of $\tilde{d}$, as

$$D = H_b^{-1}[H_b(\theta_1) + H_b(\theta_2) - H_b(\theta_1 * \theta_2)]. \quad (26)$$

It should be emphasized here that the optimal decision acts as a universal lower bound on the HD for specific schemes which assume sequential decoding. However, in the design of practical encoding/decoding algorithms, we do not consider this decision rule.

In summary, the distortion level D of the two decision rules described above is given as

$$D = \begin{cases} \min\{\theta_1, \theta_2\}, & \text{majority decision,} \\ H_b^{-1}[H_b(\theta_1) + H_b(\theta_2) - H_b(\theta_1 * \theta_2)], & \text{optimal.} \end{cases} \quad (27)$$

B. Distortion Minimization

By substituting the rate-distortion function (22) and (27) into the minimization problem (4), we have

$$\begin{aligned} &\min_{D_1, D_2} \quad D && (28)\\ &\text{s.t.}\\ &H_b[\rho * H_b^{-1}(1 - \frac{C(\gamma_2)}{r_2})] - H_b(D_1) \leq \frac{C(\gamma_1)}{r_1},\\ &H_b[\rho * H_b^{-1}(1 - \frac{C(\gamma_1)}{r_1})] - H_b(D_2) \leq \frac{C(\gamma_2)}{r_2},\\ &1 + H_b(\rho) - H_b(D_1) - H_b(D_2) \leq \frac{C(\gamma_1)}{r_1} + \frac{C(\gamma_2)}{r_2},\\ &D_i \leq \frac{1}{2}, \quad i = 1, 2,\\ &D_i \geq 0, \quad i = 1, 2. \end{aligned}$$

The reason of using the derived outer bound, not the Berger-Tung inner bound is that, the outer bound can be easily formulated as a convex optimization. The Berger-Tung inner bound includes term $D_1 * D_2$ in the binary entropy function which cannot be easily handled in the minimization. It is found that distortion $D = f(D_1, D_2)$ is monotonically increasing function on the intervals $D_i \in [0, \frac{1}{2}]$, $i = 1, 2$ for both the majority decision and optimal decision rules, and the proof is detailed in Appendix B. Furthermore, since the sequential decoding (first reconstructs $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$, then makes decision on $\mathbf{X}^n$) is applied, we first minimize the ℓ_2-norm of a vector

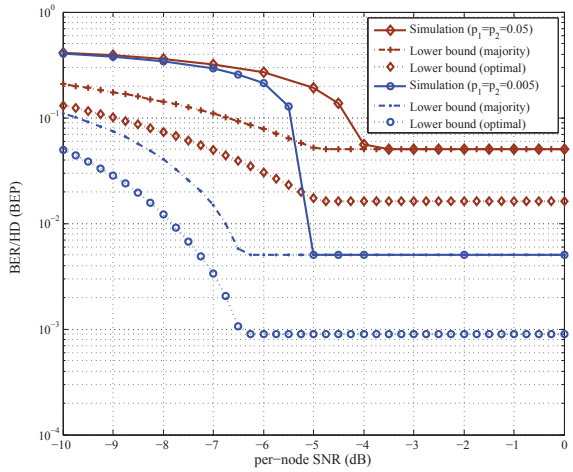

Fig. 8. Symmetric P and SNR. BPSK is used for both nodes.

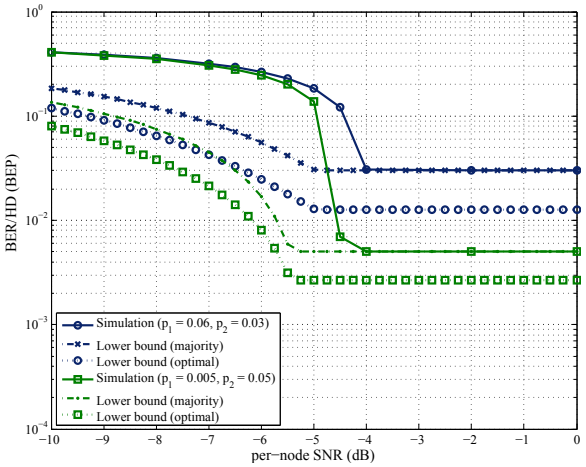

Fig. 9. Asymmetric P and symmetric SNR. BPSK is used for both nodes.

$[D_1, D_2]$ instead of directly minimizing D, as

$$\min_{D_1,D_2} \quad \|[D_1, D_2]\|_2 \tag{29}$$

s.t.

$$\begin{aligned}
&- H_b(D_1) - H_b(D_2) \le \frac{C(\gamma_1)}{r_1} + \frac{C(\gamma_2)}{r_2} - 1 - H_b(\rho),\\
&- H_b(D_1) \le \frac{C(\gamma_1)}{r_1} - H_b[\rho * H_b^{-1}(1 - \frac{C(\gamma_2)}{r_2})],\\
&- H_b(D_2) \le \frac{C(\gamma_2)}{r_2} - H_b[\rho * H_b^{-1}(1 - \frac{C(\gamma_1)}{r_1})],\\
&D_i \le \frac{1}{2}, \quad i = 1, 2,\\
&- D_i \le -0, \quad i = 1, 2,
\end{aligned}$$

to obtain the minimal values of D_1 and D_2, and then map them to D by using function $f(D_1, D_2)$.

It is easily found that the problem (29) is convex since the objective function is convex and function $-H_b(\cdot)$ is also convex. Therefore, it can be efficiently solved using convex optimization tools. Assume that the minimum values of D_1 and D_2 obtained through the convex optimization are denoted as $D_1^\star$ and $D_2^\star$, respectively. Substituting $D_1^\star$ and $D_2^\star$ into (27), the minimum distortion value $D^\star$ is then obtained through

$$D^\star = \begin{cases} \min\{\theta_1^\star, \theta_2^\star\}, & \text{majority decision,} \\ H_b^{-1}[H_b(\theta_1^\star) + H_b(\theta_2^\star) - H_b(\theta_1^\star * \theta_2^\star)], & \text{optimal,} \end{cases} \tag{30}$$

where $\theta_1^\star$ and $\theta_2^\star$ are $p_1 * D_1^\star$ and $p_2 * D_2^\star$, respectively. It should be emphasized here that the distortion D_1 or D_2 should be set to 0 in the optimization problem (28) if $\frac{C(\gamma_1)}{r_1}$ or $\frac{C(\gamma_2)}{r_2}$ is larger than or equal to 1, which is the binary entropy of the source X_1 and X_2. The reason is that a source can be reconstructed under an arbitrary small error probability if the source coding rate is larger than its entropy even in the case the helper does not exist [24].

V. BER Evaluation

A. Coding and Decoding Algorithm

We briefly review the encoding/decoding algorithm [10], [16] which is illustrated in Fig. 7. This algorithm is used to verify the theoretical HD lower bound. As illustrated in Fig. 7, each node encodes its erroneous sequence by using a serially concatenated memory-1 convolutional code and an accumulator (ACC). The encoder output sequences are then modulated and transmitted to the destination over statistically independent AWGN and block Rayleigh fading channels, where the channel gain h_i is static within each block but varies independently block-by-block. At the destination, iterative decoding process is carried out between the decoders of the convolutional code and the ACC, as well as between the two decoders of the convolutional codes through the LLR updating function f_c to modify the extrinsic LLR, according to the error probabilities p_1 and p_2.

B. Numerical Results

The lower bounds[6] on the HD for different SNR values γ_1, γ_2 are obtained through solving the convex optimization problem which we presented in Section IV. The results are shown in Figs. 8–11 for AWGN channels and Fig. 12 for block Rayleigh fading channels. The common parameters used in the simulations are

- Frame length: $n = 10000$ bits for AWGN channels and $n = 2048$ bits for block Rayleigh fading channels.
- The number of frames: 1000 for AWGN channels and 10000 for block Rayleigh fading channels.
- Interleavers: random.
- Encoder C_i: half-rate nonrecursive systematic convolutional code with generator polynomial $G = [03, 02]_8$, where $[\cdot]_8$ represents the argument is an octal number.
- Modulation: binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) with coherent detection, where channel state information is assumed to be known to the receiver. Natural mapping is used as the mapping rule in QPSK [32, Example 18.2].
- Doping ratio of ACC: 1 for BPSK and 8 for QPSK.
- Decoding algorithm for DCC_i and ACC^{-1}: log-maximum *a posteriori* (MAP).

[6]The terminology "lower bound" used here is due to the HD is calculated based on the derived outer bound, even though the approximation of the objective functions is used.

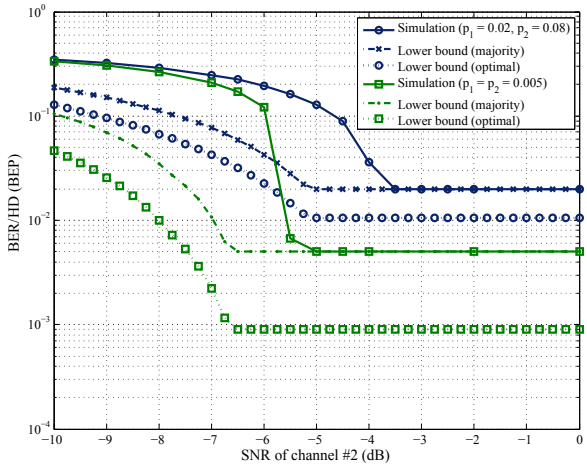

Fig. 10. Asymmetric r_1 and r_2. The coding rates r_1 and r_2 are set at $\frac{1}{4}$ and $\frac{1}{2}$, respectively. The transmit power of two nodes are the same. BPSK is used for both nodes.

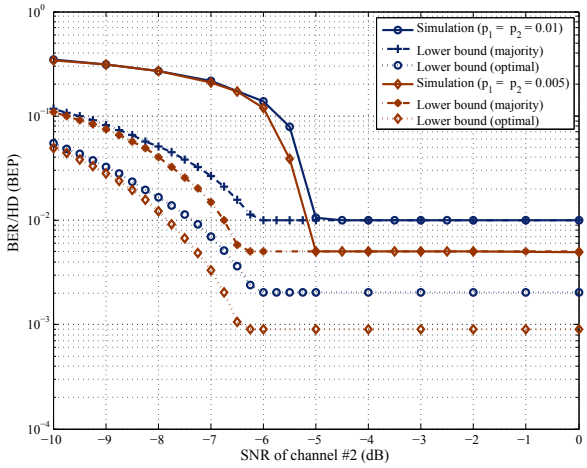

Fig. 11. Asymmetric r_1 and r_2. The coding rates r_1 and r_2 are set at 1 and $\frac{1}{2}$, respectively. The transmit power of two nodes are the same. QPSK is used for node 1 and BPSK for node 2.

- The number of iterations: 30 times.

Fig. 8 shows the error probability lower bounds and the BER versus SNR when p_1, p_2 and SNRs of the two nodes are set identically; this is referred as the symmetric case. It can be found that, the BER curves obtained by simulations and the theoretical lower bounds on the HD (or BEP) exhibit a similar tendency. Furthermore, it is clearly found that the error floor of the BER obtained by the simulation and the lower bound on the HD based on majority decision match exactly. The reason is that if the SNRs of two nodes are large enough, the distortion levels D_1 and D_2 are almost 0, which results in the error floor being determined completely by the error probabilities p_1 and p_2. A gap clearly appears between the HD lower bounds using the majority and optimal decision rules. The reason is twofold: 1) the optimality of the majority decision can not be guaranteed; 2) optimal decision is derived based on the assumption of the binary rate-distortion function without considering any loss during processing the information. To find a better decision rule than majority decision rule is left as a future study. However, it is clear that the HD lower bound deriving from the optimal decision cannot be exceeded.

The impact of the variation of the error probabilities p_1, p_2 and the coding rates r_i are evaluated in AWGN channels. Fig. 9 shows the results for asymmetric p_1 and p_2 but symmetric SNRs. When the coding rates[7] r_1 and r_2 are set as $\frac{1}{4}$ and $\frac{1}{2}$, respectively, the BER performance shown in Fig. 10 is obtained. We further consider using different modulation schemes for the nodes to achieve different rates of the channel code in Fig. 11, where QPSK is used for node 1 and BPSK for node 2. Even in these asymmetric cases, the theoretical lower bounds on the HD can still provide us with a useful reference when we evaluate the BER performance of practical systems. Furthermore, the theoretical lower bounds on the HD obtained based on our derived outer bound exhibit similar behaviors to those of the BER curves found by simulations.

[7]We simply transmit the output of ACC without doping to achieve rate $\frac{1}{4}$. No optimized design of the channel code is considered.

In both the symmetric and asymmetric cases, the threshold SNR value at which turbo cliff in the BER obtained by the simulation is around 1.5 dB larger than that observed in the theoretical lower bounds in static AWGN channels. In addition, since the lower bounds on the HD plateaus at a certain level even if the power is increased at high SNR regime, increasing the number of nodes is a proper way to improve performance in the practical deployment.

In Fig. 12, the channels between two nodes and the destination experience independent block Rayleigh fading. Therefore, the instantaneous SNRs of two nodes are different while the average SNRs of the two channels are the same. The lower bounds on the HD shown in Fig. 12 are calculated as

$$D^{\star}_{\text{fading}} = \int_0^{+\infty}\int_0^{+\infty} D^{\star}(\gamma_1,\gamma_2)\cdot \Pr(\gamma_1)\cdot \Pr(\gamma_2) d\gamma_1 d\gamma_2, \tag{31}$$

where $D^{\star}(\gamma_1,\gamma_2)$ is the result of (30), obtained for static AWGN channels. $\Pr(\gamma_i)$ is the probability density function of the SNR γ_i, which follows the Rayleigh distribution. We use Monte Carlo method to obtain the lower bounds on the average HD $D^{\star}_{\text{fading}}$ instead of theoretically calculating (31). In the Rayleigh fading case, the shape of the BER curves and the lower bounds on the HD are almost the same.

VI. Conclusion

We examined theoretically the lower bound on the HD for the binary CEO problem, where two independent nodes forward the erroneous versions of a common binary source to the destination over static AWGN and block Rayleigh fading channels. The binary CEO problem was first formulated as the binary multiterminal source coding problem, which is the core part of the binary CEO problem. The outer bound for rate-distortion region of the binary multiterminal source coding problem was then derived based on the converse proof of the bound. The relationship between the binary CEO problem and the binary multiterminal source coding problem in terms of the distortion function has been established. According to the lossy source-channel separation theorem, the lower bound on

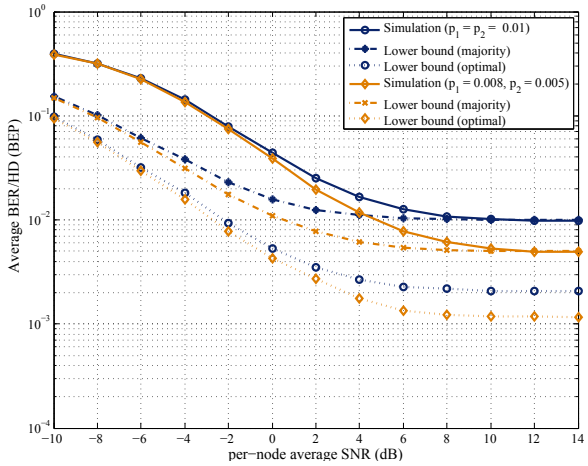

Fig. 12. Asymmetric SNR and P. BPSK is used for both nodes. Channels are assumed to suffer from block Rayleigh fading.

the HD was formulated by minimizing the distortion function subject to the inequalities between the derived outer bound and the channel capacities. The problem of obtaining the lower bound on the HD was solved in the framework of convex optimization, and the results of HD lower bounds only apply to schemes which use sequential decoding. Through a series of simulations, it has been shown that the BER curves obtained with a practical encoding/decoding algorithm is consistent with the result of the theoretical lower bounds on the HD. Even though we only solved the binary CEO problem having two nodes in this work, it shed light on fully theoretically analyzing the performance of more generic binary CEO problems such as that with an arbitrary number of nodes.

Acknowledgements

The authors wish to acknowledge the helpful discussions with Prof. Yasutada Oohama.

Appendix A
Lower bounded $H(\mathbf{X}_1^n|U_2)$

For the completeness of the paper, we briefly provide the core of [26, Corollary 4]. Let $(\mathbf{X}^n, \mathbf{Y}^n, W)$ be a triple of random variables, where $\mathbf{X}^n$ and $\mathbf{Y}^n$ take values from a set $\mathcal{X}^n = \{0,1\}^n$, and W belongs to an arbitrary discrete set $\mathcal{W}$. The joint probability mass function of this triple is given by

$$\begin{aligned}\Pr_{X,Y,W}(x,y,w) &= \Pr\{X=x, Y=y, W=w\} \\ &= \prod_{t=1}^{n} \Pr(y(t)|x(t)) \cdot \Pr_{X,W}(x,w). \quad (32)\end{aligned}$$

In other words, $\mathbf{Y}^n$ can be seen as the output of a BSC channel with crossover probability p_0 when $\mathbf{X}^n$ is the input, and W is conditionally independent of $\mathbf{Y}^n$ given $\mathbf{X}^n$. Then Corollary 4:

If $H(\mathbf{X}^n|W) \geq n \cdot v$, then $H(\mathbf{Y}^n|W) \geq n \cdot H_b[p_0 * H_b^{-1}(v)]$.

When we substitute $\mathbf{X}_1^n$, $\mathbf{X}_2^n$ and U_2 into the corollary, we have $H(\mathbf{X}_1^n|U_2) \geq nH_b(\rho * \beta)$ since $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$ is the input and the output of a BSC with crossover probability ρ, and as we assumed $H(\mathbf{X}_2^n|U_2) \geq H_b(\beta)$.

Appendix B
Monotonicity of Distortion D

Majority decision. $D = \min\{\theta_1, \theta_2\}$. Since θ_i, $i = 1, 2$ is the result of the binary convolution on p_i and D_i, θ_i is obviously increasing as D_i is increasing, when p_i is fixed.

Optimal decision. $D = H_b^{-1}[H_b(\theta_1)+H_b(\theta_2)-H_b(\theta_1*\theta_2)]$.

In this case, D is a composite function of $H_b^{-1}(\cdot)$ and $H_b(\theta_1) + H_b(\theta_2) - H_b(\theta_1 * \theta_2)$. Since the function $H_b^{-1}(\cdot)$ is monotonically increasing, we only need to prove that $g(\theta_1, \theta_2) = H_b(\theta_1)+H_b(\theta_2)-H_b(\theta_1*\theta_2)$ is also an increasing function of θ_1 and θ_2.

Assume θ_2 is fixed. The partial derivative $\frac{\partial g(\theta_1,\theta_2)}{\partial \theta_1}$ on θ_1 is

$$\frac{\partial g(\theta_1,\theta_2)}{\partial \theta_1} = \log\frac{1-\theta_1}{\theta_1} - (1-2\theta_2)\cdot\log\frac{1-\theta_1*\theta_2}{\theta_1*\theta_2}. \quad (33)$$

In order to prove that (33) is nonnegative, we should prove

$$\frac{1-\theta_1}{\theta_1} \geq \left(\frac{1-\theta_1*\theta_2}{\theta_1*\theta_2}\right)^{(1-2\theta_2)}. \quad (34)$$

The above always holds according to the monotonically increasing property of function $\log(\cdot)$. As $0 \leq \theta_i \leq \frac{1}{2}$, $i = 1, 2$ and $0 \leq \theta_1 * \theta_2 \leq \frac{1}{2}$ is assumed, the following inequalities are obtained after several steps of elementary calculation

$$\frac{1-\theta_1}{\theta_1} \geq \left(\frac{1-\theta_1}{\theta_1}\right)^{(1-2\theta_2)} \geq \left(\frac{1-\theta_1*\theta_2}{\theta_1*\theta_2}\right)^{(1-2\theta_2)}. \quad (35)$$

Therefore, it is found that (33) can not take negative values according to (35). Symmetrically, we can assume θ_1 is fixed, and show that the partial derivative $\frac{\partial g(\theta_1,\theta_2)}{\partial \theta_2}$ on θ_2 is also nonnegative. Hence, $g(\theta_1, \theta_2)$ is increasing in the dimension of θ_1 and θ_2, respectively. Based on the above two cases, it is concluded that the distortion D is increasing with respect to D_1 and D_2.

References

[1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[2] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 43, no. 5, pp. 1549–1559, Sept. 1997.

[3] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[4] J. Chen, X. Zhang, T. Berger, and S. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[5] S. Draper and G. W. Wornell, "Side information aware coding strategies for sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 966–976, Aug. 2004.

[6] H. Behroozi and M. Soleymani, "Performance of the successive coding strategy in the CEO problem," in *Proc. IEEE Global Telecommun. Conf.*, vol. 3, 28 Nov.-2 Dec. 2005, pp. 1347–1352.

[7] ——, "Optimal rate allocation in successively structured Gaussian CEO problem," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 627–632, Feb. 2009.

[8] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors I. Fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[9] J. Yuan and W. Yu, "Joint source coding, routing and power allocation in wireless sensor networks," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 886–896, June 2008.

[10] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Trans. Commun.*, vol. E95-B, no. 12, pp. 3654–3662, Dec. 2012.

[11] J. Haghighat, H. Behroozi, and D. Plant, "Iterative joint decoding for sensor networks with binary CEO model," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, July 2008, pp. 41–45.

[12] A. Razi, K. Yasami, and A. Abedi, "On minimum number of wireless sensors required for reliable binary source estimation," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Mexico, Mar. 2011, pp. 1852–1857.

[13] A. Razi and A. Abedi, "Adaptive bi-modal decoder for binary source estimation with two observers," in *Proc. Conf. Inform. Sciences Syst. (CISS)*, Princeton, NJ, USA, Mar. 2012, pp. 1–5.

[14] ——, "Convergence analysis of iterative decoding for binary CEO problem," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2944–2954, May 2014.

[15] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 555–559, Apr. 2005.

[16] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Estimation of Observation Error Probability in Wireless Sensor Networks," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1073–1076, June 2013.

[17] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," in *Proc. IEEE Inform. Theory Workshop*, Oct. 2004, pp. 315–317.

[18] M. Sartipi and F. Fekri, "Source and channel coding in wireless sensor networks using LDPC codes," in *Proc. IEEE Conf. Sensor Ad Hoc Commun. Networks*, Oct. 2004, pp. 309–316.

[19] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Wireless mesh networks allowing intra-link errors: CEO problem viewpoint," in *Proc. IEEE Int. Symp. Inform. Theory and its Applications*, Hawaii, USA, Oct. 2012, pp. 61–65.

[20] X. Zhou, M. Cheng, X. He, and T. Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7062–7071, Dec. 2014.

[21] P.-S. Lu, X. Zhou, K. Anwar, and T. Matsumoto, "Joint adaptive network-channel coding for energy-efficient multiple-access relaying," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2298–2305, June 2014.

[22] A. Gamal and Y. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[23] J.-J. Xiao and Z.-Q. Luo, "Multiterminal source-channel communication over an orthogonal multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 9, pp. 3255–3264, Sept. 2007.

[24] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. USA: John Wiley & Sons, Inc., 2006.

[25] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.

[26] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequence and applications: Part I," *IEEE Trans. Inform. Theory*, vol. 19, no. 6, pp. 769–772, Nov. 1973.

[27] S.-Y. Tung, "Multiterminal Source Coding," Ph.D. dissertation, Cornell University, 1978.

[28] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[29] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[30] ICT-619555 RESCUE Deliverable, "D1.2.1-Assessment on Feasibility, Achievability, and Limits," *Technical Report*, May 2015.

[31] V. Kostina and S. Verdu, "Fixed-length lossy compression in the finite blocklength regime," *IEEE Trans. Inform. Theory*, vol. 58, no. 6, pp. 3309–3338, June 2012.

[32] S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2004.