Title	Performance Analysis of Distortion-Acceptable Cooperative Communications in Wireless Sensor Networks for Internet of Things
Author(s)	Lin, Wensheng; Matsumoto, Tad
Citation	IEEE Sensors Journal
Issue Date	2018-11-22
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/15510
Rights	This is the author's version of the work. Copyright (C) 2018 IEEE. IEEE Sensors Journal, 2018, DOI:10.1109/JSEN.2018.2882839. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	

Performance Analysis of Distortion-Acceptable Cooperative Communications in Wireless Sensor Networks for Internet of Things

Wensheng Lin, *Student Member, IEEE*, and Tad Matsumoto, *Fellow, IEEE*

***Abstract*—This paper analyzes the performance limit of cooperative communications in wireless sensor networks (WSNs), with the aim of utilizing the analytical results to practical applications, for example, Internet of Things (IoT). The observation of a primary sensor is not necessarily to be reconstructed losslessly, as long as the system can still make correct judgements and operations. First of all, we perform the theoretical analysis to derive an inner bound on the achievable rate-distortion region for lossy communications with helpers. The numerical results precisely match the Wyner-Ziv theorem when there is only one assisting link and no rate constraint on the assisting link. Then, we present a distributed encoding and joint decoding scheme for cooperative communications in WSNs. Moreover, a series of simulations are conducted for the performance evaluation and verification. Although there is an obvious gap between the theoretical and simulation results, the performance curves show similar tendencies in terms of the signal-to-noise ratio (SNR) versus bit error rate (BER).**

***Index Terms*—Wireless sensor networks, Internet of Things, cooperative communications, rate-distortion, side information.**

I. Introduction

Wireless sensor networks (WSNs) are becoming a core part to support Internet of Things (IoT) and smart societies in the big data era [1]–[4]. Traditionally, lossless recovery of the information is needed in various communications scenarios which require high fidelity and reliability. There are already some research achievements related to lossless communications in WSNs. Zou *et al.* [5] proposed a data coding and transmission method, which can losslessly recover the original data despite the data loss occurred during transmissions, for structural health monitoring by wireless smart sensor network. In [6], Long and Xiang developed a lossless data compression algorithm based on run-length encoding and Huffman coding for energy saving in WSNs. Dedeoglu *et al.* [7] presented a distributed optimization algorithm for power allocation in lossless data gathering WSNs.

However, in the WSNs for IoT, a major task of the systems is to make a judgement followed by an operation based on the estimate of the source. Therefore, the system is still able to make correct judgement and operation, even if the source estimate is not lossless but is within a specified degree.

W. Lin is with the School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1292, Japan (email: linwest@jaist.ac.jp).

T. Matsumoto is with the School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1292, Japan and also with the Centre for Wireless Communication, University of Oulu, Oulu FI-90014, Finland (email: matumoto@jaist.ac.jp)

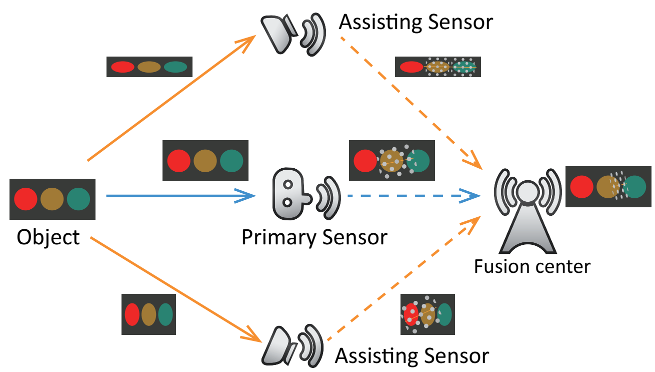


Fig. 1. A scenario of lossy communications in WSNs.

For instance, there is a primary sensor observing an object as illustrated in Fig. 1; meanwhile, a number of assisting sensors are deployed to refine the system performance. Due to the different locations and observing angles among the sensors, the observations by the assisting sensors are correlated with that of the primary sensor but not exactly the same. Then, all sensors transmit their observations to a fusion center through wireless channels, while the signals of data sequences suffer from noise. Finally, the fusion center reconstructs the observation of the primary sensor with the aid of the side information provided by the assisting sensors. Although the distortion cannot be completely eliminated in the estimate of the observation, the fusion center may be able to still make a right judgement, if the distortion is not too large. Since the estimate of the object is mainly based on the observation of the primary sensor, noiseless observation is assumed for the primary sensor which only suffers from the sensor-center transmission errors causing in the observation part.

Motivated by the realistic scenario stated above, this paper focuses on cooperative communications which allows distortions in WSNs. In order to analyze the system performance, we make our contributions on to the both theoretical analysis and performance verification through simulations in this paper. In the theoretical analysis, the distortion resulting from channel conditions can be handled by Shannon's lossy source-channel separation theorem [8], [9]. In this theorem, the sequence is encoded by lossy source coding such that the corresponding lossy coding rate *times* end-to-end coding rate is less than or equal to the channel capacity, i.e., losslessly transmitting the lossy version of the sequence via the channel, and thereby the distortion is evaluated by the rate-distortion function.

Likewise, we can start theoretical analysis from a problem of lossy multiterminal source coding, and then take the channel capacities into account for joint source-channel coding based on Shannon's lossy source-channel separation theorem.

Essentially, the theoretical model of Fig. 1 is lossy communications with helpers, i.e., the assisting sensors are the helpers from the view of the primary sensor. Thus, we begin to theoretically analyze the system performance by investigating the achievable rate-distortion region of a lossy source coding problem with helpers. So far, there are already some theoretical results related to multiterminal source coding with helpers or side information. Oohama studied a lossy source coding problem with many helpers for Gaussian sources in [10], where conditionally independent side information is assumed for a target source. For lossy source coding problem with noncausal side information only available at the decoder, Wyner and Ziv determined the corresponding rate-distortion function in [11], where the system model assumes an unconstrained full-rate helper helping one source. Ahlswede and Korner [12] characterized the rate region for estimating a source in a high fidelity with the assistance of a helper. Berger [13] and Tung [14] derived an outer bound and an inner bound of the achievable rate-distortion region if the full recoveries of the sources are not necessarily required in the multiterminal source coding problem. Han and Kobayashi studied a multiterminal source coding problem for losslessly reconstructing many sources with many helpers in [15], where an inner bound is derived by utilizing a coding scheme based on the joint typical sequence [16]. Inspired by these theoretical works, we have derived an inner bound on the achievable rate-distortion region of the lossy source coding problem with helpers for general sources. Subsequently, we determine the final distortion restricted by the channel conditions, based on Shannon's lossy source-channel separation theorem.

In order to evaluate the practical performance of lossy cooperative communications in WSNs, we present a distributed encoding and joint decoding scheme, and evaluate its bit error rate (BER) performance via computer simulations. Since the data sequences of the sensor observations are binary, the assisting data sequences can be regarded as a similar version of the primary data sequence with some bits being flipped. For binary data gathering in WSNs, We can find similar simulation system model in [17], i.e., the fusion center recursively performs soft decoding and updates log-likelihood ratio (LLR) by exchanging the mutual information among the data sequences.

The contributions of this paper are summarized as follows:

- Initially, We derive an inner bound on the achievable rate-distortion region for lossy source coding with helpers through achievability proof, where the type of distribution is not specified for the source.
- For the helper information being independent with each other, given the source, we further calculate the rate-distortion function for doubly symmetric binary source (DSBS), and extend the results to joint source-channel coding. By doing this, we analyze the theoretical performance in terms of binary distortion, or equivalently BER performance, for lossy cooperative communications in WSNs. The theoretical results are consistent to the Wyner-Ziv theorem as the special case in the sense that there is only one full-rate assisting sensor in the system.
- Finally, we perform a series of simulations to evaluate the practical performance for lossy cooperative communications in WSNs. We also discuss the reasons for the performance gap between the theoretical and simulation results. This part of the work has intuitive meaning for the future design of practical cooperative communications systems in WSNs, when lossless communication is not necessarily required.

The rest of this paper is organized as follows. Section II formulates the theoretical model as a problem of lossy communications with helpers. Section III presents an achievability proof for the rate-distortion region and derives a single-letter characterization of the inner bound. In Section IV, the rate-distortion function is further calculated for binary sources and extended to joint source-channel coding. Then, Section V evaluates the practical performance for an instance of cooperative communications systems in WSNs. Finally, Section VI concludes this work.

II. System Model

TABLE I
NOTATIONS

Notation	Definition
Uppercase letters $X, Y, \cdots$	Random variables
Lowercase letters $x, y, \cdots$	Realizations of random variables
Calligraphic letters $\mathcal{X}, \mathcal{Y}, \cdots$	Finite alphabets of variables
$\lvert \cdot \rvert$	The cardinality of a set
$\mathcal{L}$	$\{1, 2, \cdots, L\}$
$\mathcal{S}$	A subset of $\mathcal{L}$
$\mathcal{S}^c$	The complementary set of $\mathcal{S}$
S_j	The j-th element of the set $\mathcal{S}$
S_j^k	$\{S_j, S_{j+1}, \cdots, S_{k-1}, S_k\}$
t	The time index
i	The source link index
The superscript of a variable	The length of a vector
The random variable with a finite alphabet as subscript	A set of all random variables with an index in the finite alphabet, e.g., $Y_{\mathcal{S}} = \{Y_i \mid i \in \mathcal{S}\}$
$\mathcal{T}_\epsilon^{(n)}$	The set of jointly ϵ-typical n-sequences

The notations used in this paper are listed in Table I. For the purpose of simplicity, we assume that all of the sensors transmit data sequences through orthogonal channels[1]. Therefore, by Shannon's lossy source-channel separation theorem, we can consider a multiterminal source coding problem depicted in

[1]The transmission orthogonality among the links can be easily satisfied by time-division multiple access (TDMA) or orthogonal frequency-division multiple access (OFDMA). Since the link rates are constrained by signal-to-noise ratio (SNR) or signal-to-interference-plus-noise ratio (SINR) for orthogonal or non-orthogonal channels, respectively, it is not difficult to extend the results in this paper to non-orthogonal case, so far as single user detection is assumed. The use of the multiuser detection schemes is out of the scope and left as future research.

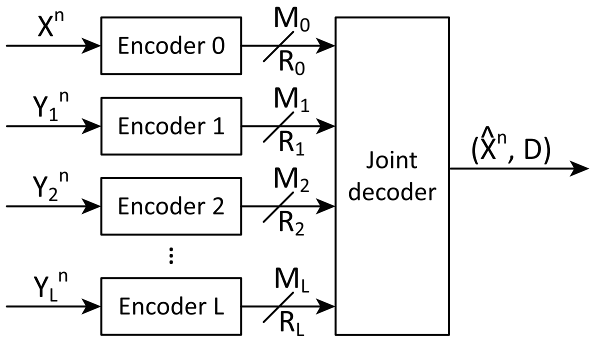


Fig. 2. The system model of lossy source coding with helpers. The primary sensor is regarded as the target source X, while the assisting sensors are regarded as many helpers Y_i.

Fig. 2 for the first step. The primary sensor is regarded as the target source X, while the assisting sensors are regarded as many helpers Y_i. Although the number of the helpers may be limited by space and channel resources in practical scenario, we do not constrain it in theoretical analysis. In total, there are $(L+1)$ memoryless sources $(X, Y_1, \cdots, Y_L)$, with X and Y_i taking values from corresponding finite alphabets $\mathcal{X}$ and $\mathcal{Y}_i$ at each time slot, respectively. $x^n = \{x(t)\}_{t=1}^n$ and $y_i^n = \{y_i(t)\}_{t=1}^n$ indicate independent and identically distributed (i.i.d.) sequences from the sources X and Y_i, respectively. Then, the sequences x^n and y_i^n are observed and transmitted to a common receiver, i.e., the fusion center, after compression by encoder 0 and encoder i, respectively. Due to some restrictions in practice, e.g., the deployment of encoders is distributed and located at different places, the observed sequences x^n and y_i^n have to be encoded into codewords separately. The encoders compress the sequences x^n and y_i^n at rates R_0 and R_i, respectively, by assigning an index to each sequence according to the following mapping rules:

$$\varphi_0 : \mathcal{X}^n \mapsto \mathcal{M}_0 = \{1, 2, \cdots, 2^{nR_0}\}, \tag{1}$$

$$\varphi_i : \mathcal{Y}_i^n \mapsto \mathcal{M}_i = \{1, 2, \cdots, 2^{nR_i}\}, \text{ for } i \in \mathcal{L}. \tag{2}$$

Without loss of generality, we assume that 2^{nR_0} and 2^{nR_i} are integer numbers.

After receiving all of the encoding outputs $\varphi_0(x^n), \varphi_1(y_1^n), \cdots, \varphi_L(y_L^n)$, a joint decoder constructs the estimate $\hat{x}^n$ of x^n by *jointly* utilizing the received codewords in contrast to the *distributed compression* at the encoders. The reconstruction process is expressed by the mapping as:

$$\psi : \mathcal{M}_0 \times \mathcal{M}_1 \times \cdots \times \mathcal{M}_L \mapsto \mathcal{X}^n. \tag{3}$$

Generally, distortion happens when the estimate $\hat{x}^n$ does not fully contain the information of x^n. The distortion measure $d : \mathcal{X} \times \mathcal{X} \mapsto [0, \infty)$ is defined to describe the degree of distortion between x and $\hat{x}$. For the entire sequence, the average distortion between the sequences x^n and $\hat{x}^n$ is defined as

$$d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{t=1}^{n} d\left(x(t), \hat{x}(t)\right). \tag{4}$$

For given distortion requirement D, the achievable rate-distortion region $\mathcal{R}(D)$, consisting of all achievable rate tuple of $(R_0, R_{\mathcal{L}})$, is defined as

$$\mathcal{R}(D) = \{(R_0, R_{\mathcal{L}}) : (R_0, R_{\mathcal{L}}) \text{ is admissible such that } \lim_{n\to\infty} \mathrm{E}\left(d(x^n, \hat{x}^n)\right) \le D + \epsilon, \text{ for any } \epsilon > 0\}. \tag{5}$$

Since the channels are assumed to be orthogonal, according to Shannon's lossy source-channel separation theorem, the link rates are constrained by:

$$\begin{cases} R_0(D) \cdot r_0 \le C(\gamma_0), \\ \quad R_i \cdot r_i \le C(\gamma_i), \text{ for } i \in \mathcal{L}, \end{cases} \tag{6}$$

where r_0 and r_i denote the end-to-end coding rates; moreover, $C(\gamma_0)$ and $C(\gamma_i)$ represent the Shannon capacity using Gaussian codebook with γ being the SNR of the wireless channel.

III. Lossy Source Coding with Helpers

Initially, we derive a single-letter characterization of an inner bound on the achievable rate-distortion region for lossy source coding with helpers.

Proposition 1: Let $(X, Y_{\mathcal{L}})$ be a $(L+1)$-component discrete memoryless source and $d(x, \hat{x})$ be distortion measure. A rate tuple $(R_0, R_{\mathcal{L}})$ is achievable with distortion requirement D for distributed lossy source coding with more-than-one helpers if

$$R_0 > I(X; U | V_{\mathcal{L}}), \tag{7}$$

$$\sum_{i \in \mathcal{S}} R_i > I(Y_{\mathcal{S}}; V_{\mathcal{S}} | V_{\mathcal{S}^c}), \tag{8}$$

for some conditional probability mass function (PMF) $p(u|x) \cdot \prod_{i=1}^{L} p(v_i|y_i)$ and function $\hat{x}(u, v_{\mathcal{L}})$ such that $\mathrm{E}(d(X, \hat{X})) \le D$, with $U \to X \to Y_i \to V_i$ and $V_i \to Y_i \to X \to Y_j \to V_j$ forming Markov chains for $i, j \in \mathcal{L}$ and $i \ne j$.

Proof of Proposition 1: We use a $(L+1)$-dimension distributed compress-bin scheme for lossy source coding, and analyze the expected distortion of this scheme with respect to rate constraints. In the following, we assume that $\epsilon_1 < \epsilon_2 < \epsilon_3 < \epsilon$.

Codebook generation. Fix a conditional PMF $p(u|x) \cdot \prod_{i=1}^{L} p(v_i|y_i)$ and a function $\hat{x}(u, v_{\mathcal{L}})$ such that $\mathrm{E}(d(X, \hat{X})) \le D/(1+\epsilon)$. Let $\tilde{R}_0 \ge R_0$ and $\tilde{R}_i \ge R_i$ for $i \in \mathcal{L}$. Randomly and independently generate $2^{n\tilde{R}_0}$ sequences $u^n(k_0) \sim \prod_{t=1}^{n} p_U(u), k_0 \in \mathcal{K}_0 = \{1, 2, \cdots, 2^{n\tilde{R}_0}\}$. Similarly, for $i \in \mathcal{L}$, randomly and independently generate $2^{n\tilde{R}_i}$ sequences $v_i^n(k_i) \sim \prod_{t=1}^{n} p_{V_i}(v_i), k_i \in \mathcal{K}_i = \{1, 2, \cdots, 2^{n\tilde{R}_i}\}$. Partition the set of indices $k_0 \in \mathcal{K}_0$ into equal-size bins $\mathcal{B}_0(m_0) = \{(m_0 - 1)2^{n(\tilde{R}_0 - R_0)} + 1, \cdots, m_0 2^{n(\tilde{R}_0 - R_0)}\}$ for $m_0 \in \mathcal{M}_0$, and also partition the set of indices $k_i \in \mathcal{K}_i$ into equal-size bins $\mathcal{B}_i(m_i) = \{(m_i - 1)2^{n(\tilde{R}_i - R_i)} + 1, \cdots, m_i 2^{n(\tilde{R}_i - R_i)}\}$ for $m_i \in \mathcal{M}_i, i \in \mathcal{L}$. This codebook structure is utilized in the encoders and the decoder.

Encoding. Upon observing x^n, encoder 0 finds an index $k_0 \in \mathcal{K}_0$ such that $(u^n(k_0), x^n) \in \mathcal{T}_{\epsilon_1}^{(n)}$. If there is more than one such index k_0, encoder 0 selects one of them uniformly at random. If there is no such index k_0, encoder 0 selects an index from $\mathcal{K}_0$ uniformly at random. Similarly, for $i \in \mathcal{L}$, encoder i finds an index $k_i \in \mathcal{K}_i$ such that $(v_i^n(k_i), y_i^n) \in \mathcal{T}_{\epsilon_1}^{(n)}$. If there is more than one such index k_i, encoder i selects one of them uniformly at random. If there is no such index k_i, encoder i selects an index from $\mathcal{K}_i$ uniformly at random. Then, encoder 0 and encoder i send the indices m_0 and m_i such that $k_0 \in \mathcal{B}_0(m_0)$ and $k_i \in \mathcal{B}_i(m_i)$, respectively.

Decoding. The decoder finds the unique index tuple $(\hat{k}_0, \hat{k}_{\mathcal{L}}) \in \mathcal{B}_0(m_0) \times \mathcal{B}_1(m_1) \times \cdots \times \mathcal{B}_L(m_L)$ such that $(u^n(\hat{k}_0), v_1^n(\hat{k}_1), \cdots, v_L^n(\hat{k}_L)) \in \mathcal{T}_\epsilon^{(n)}$. If there is such a unique index tuple $(\hat{k}_0, \hat{k}_{\mathcal{L}})$, the reconstruction is computed bit by bit as $\hat{x}_t(u_t(\hat{k}_0), v_{1,t}(\hat{k}_1), \cdots, v_{L,t}(\hat{k}_L))$; otherwise, $\hat{x}^n$ is set to arbitrary sequence in $\mathcal{X}^n$.

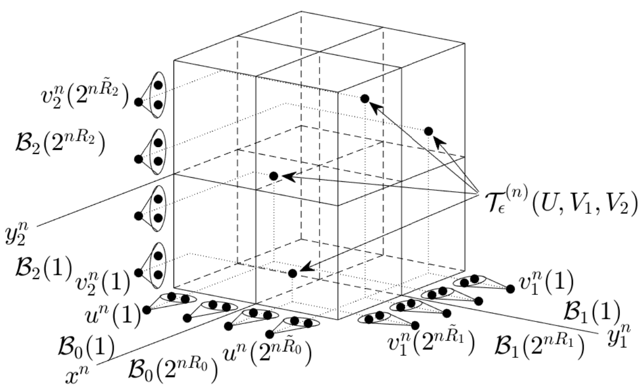


Fig. 3. An example of the distributed compress-bin scheme with $L = 2$.

An example of the distributed compress-bin scheme with $L = 2$ is depicted in Fig. 3. Now, we analyze the expected distortion of the distributed compress-bin scheme. Let $(K_0, K_{\mathcal{L}})$ denote the index tuple for the chosen $(U^n, V_{\mathcal{L}}^n)$ tuple, $(M_0, M_{\mathcal{L}})$ be the tuple of corresponding bin indices, and $(\hat{K}_0, \hat{K}_{\mathcal{L}})$ be the tuple of decoded indices. Define the "error" event

$$\begin{aligned}\mathcal{E} = \{&(U^n(\hat{K}_0), V_1^n(\hat{K}_1), \cdots, V_L^n(\hat{K}_L), X^n, Y_1^n, \cdots, Y_L^n) \\ &\notin \mathcal{T}_\epsilon^{(n)}\},\end{aligned} \tag{9}$$

and consider the following events:

$$\mathcal{E}_1 = \{(U^n(k_0), X^n) \notin \mathcal{T}_{\epsilon_1}^{(n)} \text{ for all } k_0 \in \mathcal{K}_0\}, \tag{10}$$

$$\mathcal{E}_2 = \{(V_i^n(k_i), Y_i^n) \notin \mathcal{T}_{\epsilon_1}^{(n)} \text{ for all } k_i \in \mathcal{K}_i, i \in \mathcal{L}\}, \tag{11}$$

$$\mathcal{E}_3 = \{(U^n(K_0), X^n, Y_1^n) \notin \mathcal{T}_{\epsilon_2}^{(n)}\}, \tag{12}$$

$$\mathcal{E}_4 = \{(U^n(K_0), X^n, V_1^n(K_1), Y_1^n) \notin \mathcal{T}_{\epsilon_3}^{(n)}\}, \tag{13}$$

$$\begin{aligned}\mathcal{E}_5 = \{&(U^n(K_0), X^n, V_1^n(K_1), Y_1^n, \cdots, V_L^n(K_L), Y_L^n) \\ &\notin \mathcal{T}_\epsilon^{(n)}\},\end{aligned} \tag{14}$$

$$\begin{aligned}\mathcal{E}_6 = \{&(V_1^n(\tilde{k}_1), \cdots, V_L^n(\tilde{k}_L)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some} \\ &\tilde{k}_{\mathcal{L}} \in \mathcal{B}_1(M_1) \times \cdots \times \mathcal{B}_L(M_L), \tilde{k}_{\mathcal{L}} \neq K_{\mathcal{L}}\},\end{aligned} \tag{15}$$

$$\begin{aligned}\mathcal{E}_7 = \{&(U^n(\tilde{k}_0), V_1^n(K_1), \cdots, V_L^n(K_L)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some} \\ &\tilde{k}_0 \in \mathcal{B}_0(M_0), \tilde{k}_0 \neq K_0\}.\end{aligned} \tag{16}$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ represent encoding error events in encoder 0 and encoder i for $i \in \mathcal{L}$, respectively. $\mathcal{E}_5$ occurs if joint typicality decoding fails, with $\mathcal{E}_3$ and $\mathcal{E}_4$ being its sub events. $\mathcal{E}_6$ and $\mathcal{E}_7$ mean that there are more than one decoding result, and hence a decoding error event occurs. Notice that the "error" event occurs only if $(U^n(K_0), V_1^n(K_1), \cdots, V_L^n(K_L), X^n, Y_1^n, \cdots, Y_L^n) \notin \mathcal{T}_\epsilon^{(n)}$ or $(\tilde{k}_0, \tilde{k}_{\mathcal{L}}) \neq (K_0, K_{\mathcal{L}})$. By the union of the events bound, we have

$$\begin{aligned}\mathrm{P}(\mathcal{E}) \le\ &\mathrm{P}(\mathcal{E}_1) + \mathrm{P}(\mathcal{E}_2) + \mathrm{P}(\mathcal{E}_1^c \cap \mathcal{E}_3) + \mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4) \\ &+ \mathrm{P}(\mathcal{E}_4^c \cap \mathcal{E}_5) + \mathrm{P}(\mathcal{E}_6) + \mathrm{P}(\mathcal{E}_7).\end{aligned} \tag{17}$$

We bound each term as follows. First, by the covering lemma [18], $\mathrm{P}(\mathcal{E}_1)$ tends to zero as $n \to \infty$ if

$$\tilde{R}_0 > I(X; U) + \delta(\epsilon_1), \tag{18}$$

and $\mathrm{P}(\mathcal{E}_2)$ tends to zero as $n \to \infty$ if

$$\tilde{R}_i > I(Y_i; V_i) + \delta(\epsilon_1). \tag{19}$$

Since $\mathcal{E}_1^c = \{(U^n(K_0), X^n) \in \mathcal{T}_{\epsilon_1}^{(n)}\}$, $Y_1^n | \{U^n(K_0) = u^n, X^n = x^n\} \sim \prod_{t=1}^n p_{Y_1|X}(y_{1,t}|x_t)$. By the conditional typicality lemma [18], $\mathrm{P}(\mathcal{E}_1^c \cap \mathcal{E}_3)$ approaches zero as $n \to \infty$.

To bound $\mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4)$, let $(u^n, x^n, y_1^n) \in \mathcal{T}_{\epsilon_2}^{(n)}(U, X, Y_1)$, and consider

$$\begin{aligned}&\mathrm{P}\{V_1^n(K_1) = v_1^n | U^n(K_0) = u^n, X^n = x^n, Y_1^n = y_1^n\} \\ &= \mathrm{P}\{V_1^n(K_1) = v_1^n | Y_1^n = y_1^n\} \\ &= p(v_1^n | y_1^n).\end{aligned} \tag{20}$$

First, notice that by the covering lemma, $\mathrm{P}\{V_1^n(K_1) \in \mathcal{T}_{\epsilon_2}^{(n)}(V_1|y_1^n) | Y_1^n = y_1^n\}$ converges to 1 as $n \to \infty$, i.e., $p(v_1^n|y_1^n)$ satisfies the first condition of the Markov lemma [18]. Then, similar to the proof of the Berger-Tung inner bound, shown in Lemma 12.3 in [18], $p(v_1^n|y_1^n)$ also satisfies the second condition of the Markov lemma. Hence, according to the Markov lemma, we have

$$\begin{aligned}\lim_{n\to\infty} &\mathrm{P}\{(u^n, x^n, y_1^n, V_1^n(K_1)) \in \mathcal{T}_{\epsilon_3}^{(n)} | U^n(K_0) = u^n, \\ &\quad X^n = x^n, Y_1^n = y_1^n\} \\ &= 1,\end{aligned} \tag{21}$$

if $(u^n, x^n, y_1^n) \in \mathcal{T}_{\epsilon_2}^{(n)}(U, X, Y_1)$ and $\epsilon_2 < \epsilon_3$ is sufficiently small. Therefore, $\mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4)$ tends to zero as $n \to \infty$. By recursively utilizing the similar derivation for bounding $\mathrm{P}(\mathcal{E}_3)$ and $\mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4)$, we can obtain that $\mathrm{P}(\mathcal{E}_4^c \cap \mathcal{E}_5)$ tends to zero as $n \to \infty$.

To bound $\mathrm{P}(\mathcal{E}_6)$, we introduce the following two lemmas:

Lemma 1 (joint typicality lemma for multiple random variables): Let $(V_{\mathcal{S}}, V_{\mathcal{S}^c}) \sim p(v_{\mathcal{S}}, v_{\mathcal{S}^c})$. If $\tilde{v}_i^n \sim \prod_{t=1}^n p_{V_i}(v_{i,t})$ for $i \in \mathcal{S}$, and $\tilde{v}_i^n$ is an arbitrary random sequence for $i \in \mathcal{S}^c$, then

$$\begin{aligned}&\mathrm{P}\{(\tilde{V}_{\mathcal{S}}^n, \tilde{v}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}, V_{\mathcal{S}^c})\} \\ &\le \mathrm{pow}\left(2, -n\left[\sum_{j=2}^{|\mathcal{S}|} I(V_{\mathcal{S}_1^{j-1}}; V_{\mathcal{S}_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) \right.\right. \\ &\qquad \left.\left. - \delta(\epsilon)\right]\right),\end{aligned} \tag{22}$$

where $\mathrm{pow}(a, b) = a^b$.

Lemma 2 (mutual packing lemma for multiple random variables): Let $(V_{\mathcal{S}}, V_{\mathcal{S}^c}) \sim p(v_{\mathcal{S}}, v_{\mathcal{S}^c})$. For $i \in \mathcal{S}$, let $V_i^n(k_i) \sim \prod_{t=1}^n p_{V_i}(v_{i,t})$, $k_i \in \mathcal{K}_i = \{1, 2, \cdots, 2^{nr_i}\}$. For $i \in \mathcal{S}^c$, let $\tilde{V}_i^n$ be an arbitrarily distributed random sequence.

Assume that $(V_i^n(k_i) : i \in \mathcal{S}, k_i \in \mathcal{K}_i)$ and $(\tilde{V}_i^n : i \in \mathcal{S}^c)$ are independent of each other. Then, $\delta(\epsilon)$ exists that tends to zero as $\epsilon \to 0$ such that

$$\begin{aligned}\lim_{n\to\infty} \mathrm{P}\{(V_{S_1}^n(k_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(k_{S_{|\mathcal{S}|}}), \tilde{V}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some}\\ k_i \in \mathcal{K}_i, i \in \mathcal{S}\}\\ = 0, \end{aligned} \tag{23}$$

if

$$\sum_{i\in\mathcal{S}} r_i < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon). \tag{24}$$

The proofs of *Lemma 1* and *Lemma 2* are provided in Appendix A and Appendix B, respectively.

If $\tilde{R}_i = R_i$ for $i \in \mathcal{S}^c$, notice that (19) becomes

$$R_i > I(Y_i; V_i) + \delta(\epsilon_1), \tag{25}$$

and hence R_i is already large enough for link i. Moreover, since $\tilde{R}_i - R_i = 0$, there is only one index in $\mathcal{B}_i$ for $i \in \mathcal{S}^c$. Hence, $\tilde{k}_i = K_i$ for $\tilde{k}_i \in \mathcal{B}_i$. Then, $\mathcal{E}_6$ can be simplified as

$$\begin{aligned}\mathcal{E}_6 = \{&(V_{S_1}^n(\tilde{k}_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(\tilde{k}_{S_{|\mathcal{S}|}}), V_{S_1^c}^n(K_{S_1^c}), \cdots,\\ &V_{S_{|\mathcal{S}^c|}^c}^n(K_{S_{|\mathcal{S}^c|}^c})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{k}_{\mathcal{S}} \neq K_{\mathcal{S}},\\ &\tilde{k}_{\mathcal{S}} \in \mathcal{B}_{S_1}(M_{S_1}) \times \cdots \times \mathcal{B}_{S_{|\mathcal{S}|}}(M_{S_{|\mathcal{S}|}})\}. \end{aligned}\tag{26}$$

Following a similar argument as Lemma 11.1 in [18] in the proof of the Wyner-Ziv theorem, we have

$$\begin{aligned}\mathrm{P}(\mathcal{E}_6) \leq \mathrm{P}\{&(V_{S_1}^n(\tilde{k}_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(\tilde{k}_{S_{|\mathcal{S}|}}), V_{S_1^c}^n(K_{S_1^c}), \cdots,\\ &V_{S_{|\mathcal{S}^c|}^c}^n(K_{S_{|\mathcal{S}^c|}^c})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some}\\ &\tilde{k}_{\mathcal{S}} \in \mathcal{B}_{S_1}(1) \times \cdots \times \mathcal{B}_{S_{|\mathcal{S}|}}(1)\}, \end{aligned}\tag{27}$$

$$\begin{aligned}\mathrm{P}(\mathcal{E}_7) \leq \mathrm{P}\{&(U^n(\tilde{k}_0), V_1^n(K_1), \cdots, V_L^n(K_L)) \in \mathcal{T}_\epsilon^{(n)}\\ &\text{for some } \tilde{k}_0 \in \mathcal{B}_0(1)\}. \end{aligned}\tag{28}$$

According to *Lemma 2* and the packing lemma [18], $\mathrm{P}(\mathcal{E}_6)$ and $\mathrm{P}(\mathcal{E}_7)$ tend to zero as $n \to \infty$, respectively, if

$$\begin{aligned}\sum_{i\in\mathcal{S}}(\tilde{R}_i - R_i) < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c})\\ - \delta(\epsilon), \end{aligned}\tag{29}$$

$$\tilde{R}_0 - R_0 < I(U; V_{\mathcal{L}}) - \delta(\epsilon). \tag{30}$$

By combining (18), (19), (29) and (30), we have shown that $\mathrm{P}(\mathcal{E})$ tends to zero as $n \to \infty$ if

$$R_0 > I(X; U) + \delta(\epsilon_1) - I(U; V_{\mathcal{L}}) + \delta(\epsilon), \tag{31}$$

$$\begin{aligned}\sum_{i\in\mathcal{S}} R_i > \sum_{i\in\mathcal{S}}[I(Y_i; V_i) + \delta(\epsilon_1)] - \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\\ - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta(\epsilon). \end{aligned}\tag{32}$$

We can further calculate (31) as

$$\begin{aligned}R_0 &> I(X; U) + \delta(\epsilon_1) - I(U; V_{\mathcal{L}}) + \delta(\epsilon),\\ &= I(X, V_{\mathcal{L}}; U) - I(U; V_{\mathcal{L}}) + \delta'(\epsilon) &(33)\\ &= I(X; U|V_{\mathcal{L}}) + \delta'(\epsilon), &(34)\end{aligned}$$

where (33) follows since $V_{\mathcal{L}} \to Y_{\mathcal{L}} \to X \to U$ forms a Markov chain, and $\delta'(\epsilon) = \delta(\epsilon_1) + \delta(\epsilon)$. (32) can further be reduced to:

$$\begin{aligned}\sum_{i\in\mathcal{S}} R_i &> \sum_{i\in\mathcal{S}}[I(Y_i; V_i) + \delta(\epsilon_1)] - \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\\ &\quad - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta(\epsilon)\\ &= I(Y_{S_1}; V_{S_1}) + I(Y_{S_2}; V_{S_2}) - I(V_{S_1}; V_{S_2})\\ &\quad + \sum_{j=3}^{|\mathcal{S}|}[I(Y_{S_j}; V_{S_j}) - I(V_{S_1^{j-1}}; V_{S_j})] - I(V_{\mathcal{S}}; V_{\mathcal{S}^c})\\ &\quad + \delta'(\epsilon), \end{aligned}\tag{35}$$

where $\delta'(\epsilon) = |\mathcal{S}| \cdot \delta(\epsilon_1) + \delta(\epsilon)$. Consider

$$\begin{aligned}&I(Y_{S_1^{j-1}}; V_{S_1^{j-1}}) + I(Y_{S_j}; V_{S_j}) - I(V_{S_1^{j-1}}; V_{S_j})\\ &= I(Y_{S_1^{j-1}}; V_{S_1^{j-1}}) + I(Y_{S_j}; V_{S_j}) - I(V_{S_1^{j-1}}; V_{S_j})\\ &\quad + H(Y_{S_1^{j-1}}|V_{S_1^{j-1}}, Y_{S_j}) - H(Y_{S_1^{j-1}}|V_{S_1^{j-1}}, Y_{S_j})\\ &= I(Y_{S_1^{j-1}}; V_{S_1^{j-1}}) + I(Y_{S_j}; V_{S_j}) - I(V_{S_1^{j-1}}; V_{S_j})\\ &\quad + H(Y_{S_1^{j-1}}|V_{S_1^{j-1}}, Y_{S_j})\\ &\quad - H(Y_{S_1^{j-1}}|V_{S_1^{j-1}}, Y_{S_j}, V_{S_j}) &(36)\\ &= I(Y_{S_1^{j-1}}; V_{S_1^{j-1}}) + I(Y_{S_j}; V_{S_j}) - I(V_{S_1^{j-1}}; V_{S_j})\\ &\quad + I(Y_{S_1^{j-1}}; V_{S_j}|V_{S_1^{j-1}}, Y_{S_j})\\ &= I(Y_{S_1^{j-1}}, Y_{S_j}; V_{S_1^{j-1}}) + I(Y_{S_j}, V_{S_1^{j-1}}; V_{S_j})\\ &\quad - I(V_{S_1^{j-1}}; V_{S_j}) + I(Y_{S_1^{j-1}}; V_{S_j}|V_{S_1^{j-1}}, Y_{S_j}) &(37)\\ &= I(Y_{S_1^j}; V_{S_1^{j-1}}) + I(Y_{S_j}; V_{S_j}|V_{S_1^{j-1}})\\ &\quad + I(Y_{S_1^{j-1}}; V_{S_j}|V_{S_1^{j-1}}, Y_{S_j})\\ &= I(Y_{S_1^j}; V_{S_1^{j-1}}) + I(Y_{S_1^{j-1}}, Y_{S_j}; V_{S_j}|V_{S_1^{j-1}})\\ &= I(Y_{S_1^j}; V_{S_1^{j-1}}) + I(Y_{S_1^j}; V_{S_j}|V_{S_1^{j-1}})\\ &= I(Y_{S_1^j}; V_{S_1^{j-1}}, V_{S_j})\\ &= I(Y_{S_1^j}; V_{S_1^j}), &(38)\end{aligned}$$

where (36) follows the fact that V_{S_j} is a function of Y_{S_j}, and (37) follows that $V_{S_1^{j-1}} \to Y_{S_1^{j-1}} \to Y_{S_j} \to V_{S_j}$ forms a Markov chain. By substituting (38) into (35), we have

$$\begin{aligned}\sum_{i\in\mathcal{S}} R_i &> I(Y_{S_1^2}; V_{S_1^2}) + \sum_{j=3}^{|\mathcal{S}|}[I(Y_{S_j}; V_{S_j})\\ &\quad - I(V_{S_1^{j-1}}; V_{S_j})] - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta'(\epsilon)\\ &= I(Y_{S_1^{|\mathcal{S}|}}; V_{S_1^{|\mathcal{S}|}}) - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta'(\epsilon)\\ &= I(Y_{\mathcal{S}}; V_{\mathcal{S}}) - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta'(\epsilon)\\ &= I(Y_{\mathcal{S}}, V_{\mathcal{S}^c}; V_{\mathcal{S}}) - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta'(\epsilon) &(39)\\ &= I(Y_{\mathcal{S}}; V_{\mathcal{S}}|V_{\mathcal{S}^c}) + \delta'(\epsilon), &(40)\end{aligned}$$

where (39) follows the fact that $V_{\mathcal{S}^c} \to Y_{\mathcal{S}^c} \to Y_{\mathcal{S}} \to V_{\mathcal{S}}$ forms a Markov chain.

Notice that $(U^n(K_0), V_1^n(K_1), \cdots, V_L^n(K_L), X^n, Y_1^n, \cdots, Y_L^n) \in \mathcal{T}_\epsilon^{(n)}$, when there is no "error". Therefore, by the law of total expectation and the typical average lemma, the asymptotic distortion, averaged over the random codebook and encoding, is upper bounded as

$$\begin{aligned}&\lim_{n\to\infty}\sup \mathrm{E}(d(X^n,\hat{X}^n))\\ &\le \lim_{n\to\infty}\sup \Big[d_{\max}\cdot \mathrm{P}(\mathcal{E})\\ &\qquad\qquad + (1+\epsilon)\cdot \mathrm{E}(d(X,\hat{X}))\cdot \mathrm{P}(\mathcal{E}^c)\Big] \quad (41)\\ &\le D, \quad (42)\end{aligned}$$

if the inequalities in (34) and (40) are satisfied. Finally, from the continuity of mutual information and taking $\epsilon \to 0$, we complete the proof of *Proposition 1*.

IV. Rate-Distortion Analysis for Binary Sources

Here, we analyze the achievable rate-distortion region for binary sources. Consider a DSBS(p_i) (X, Y_i) with $X \sim \mathrm{Bern}(0.5)$ and $Y_i \sim \mathrm{Bern}(0.5)$ for $i \in \mathcal{L}$, which follows a joint PMF $p_{X,Y_i}(x, y_i) = \Pr\{X = x, Y_i = y_i\}$ given by

$$p_{X,Y_i}(x,y_i) = \begin{cases} \frac{1}{2}p_i, & \text{if } x \neq y_i,\\ \frac{1}{2}(1-p_i), & \text{otherwise,}\end{cases} \quad (43)$$

where the correlation parameter $p_i = \Pr\{x \neq y_i\}$, $p_i \in [0, \frac{1}{2}]$. Equivalently, Y_i can be considered to be the output of a binary symmetric channel (BSC) with input X and crossover probability p_i and vice versa. The distortion measure is set as the Hamming distortion measure for binary sources, i.e.,

$$d(x,\hat{x}) = \begin{cases} 1, & \text{if } x \neq \hat{x},\\ 0, & \text{if } x = \hat{x}.\end{cases} \quad (44)$$

A. Achievable Rate-Distortion Region for Binary Sources

Now, we calculate the constraints of the achievable rate-distortion region for DSBS. First, consider

$$\begin{aligned}R_0(D) &> I(X;U|V_{\mathcal{L}})\\ &= H(U|V_{\mathcal{L}}) - H(U|V_{\mathcal{L}},X)\\ &= H(U|V_{\mathcal{L}}) - H(U|X) \quad (45)\\ &= H(U,V_{\mathcal{L}}) - H(V_{\mathcal{L}}) - h(D), \quad (46)\end{aligned}$$

where (45) follows since $V_{\mathcal{L}} \to X \to U$ forms a Markov chain, and $h(\cdot)$ denotes the binary entropy function. In order to further calculate (46), we introduce a joint entropy function for the correlated binary sources with a set of crossover probabilities $\{\mathcal{P}\}$:

Definition 1: According to [19], given a set of crossover probabilities $\{\mathcal{P}\}$ with a common source X, the joint entropy $f(\cdot)$ of the outputs from independent BSCs is calculated as

$$f(\{\mathcal{P}\}) = -\sum_{j=1}^{2^{|\mathcal{P}|}} q_j \log_2(q_j), \quad (47)$$

where

$$q_j = 0.5\left(\prod_{k\in\mathcal{A}_i} p_k \prod_{k'\in\mathcal{A}_i^c} \bar{p}_{k'} + \prod_{k\in\mathcal{A}_i} \bar{p}_k \prod_{k'\in\mathcal{A}_i^c} p_{k'}\right), \quad (48)$$

with $\bar{p} = 1 - p$ and $\mathcal{A}_i \subseteq \{1, 2, \cdots, |\mathcal{P}|\}$.

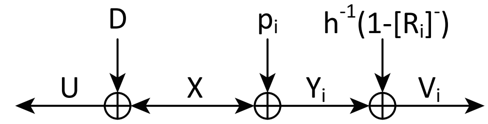


Fig. 4. The test channels for binary sources, where $[R_i]^- = \min\{1, R_i\}$, and $h^{-1}(\cdot)$ denotes the inverse function of $h(\cdot)$.

Since $V_i \to Y_i \to X \to Y_j \to V_j$ forms a Markov chain for $i \neq j$, i.e., Y_i are independent to each other if X is given, we can obtain the test channel shown in Fig. 4. Then, by *Definition 1*, we can calculate (46) as

$$R_0(D) > f(\{D, \alpha_{\mathcal{L}}\}) - f(\{\alpha_{\mathcal{L}}\}) - h(D), \quad (49)$$

where $\alpha_i = p_i * h^{-1}(1 - [R_i]^-)$, and the operation $*$ denotes the binary convolution process, i.e., $a * b = a(1-b) + b(1-a)$. Notice that $D = 0.5$ if $R_0 = 0$ according to (49). However, it is obvious that by decoding only with the compressed side information $V_{\mathcal{L}}$, $\hat{X}$ still can achieve the distortion

$$\begin{aligned}D' &= h^{-1}[H(X|V_{\mathcal{L}})]\\ &= h^{-1}[H(X,V_{\mathcal{L}}) - H(V_{\mathcal{L}})]\\ &= h^{-1}[f(\{0,\alpha_{\mathcal{L}}\}) - f(\{\alpha_{\mathcal{L}}\})], \quad (50)\end{aligned}$$

where (50) holds since X can be regarded as the output of a BSC with itself as input and the crossover probability $p_0 = 0$. Therefore, the optimal performance can be achieved by time sharing between rate-distortion coding and zero-rate decoding only with the compressed side information. Consequently, we can obtain the rate-distortion function for DSBS, as

$$R_0(D) = \begin{cases} g(D), & \text{for } 0 \le D \le D_c,\\ (D - D')g'(D_c), & \text{for } D_c < D \le D',\\ 0, & \text{for } D' < D,\end{cases} \quad (51)$$

where $g(D) = f(\{D, \alpha_{\mathcal{L}}\}) - f(\{\alpha_{\mathcal{L}}\}) - h(D)$ with $g'(D)$ being the derivative of $g(D)$, and D_c is the solution to the equation $g(D_c) = (D_c - D')g'(D_c)$.

Finally, we extend the above results of multiterminal source coding into joint source-channel coding based on Shannon's lossy source-channel separation theorem. By combining (6) and (51), we have

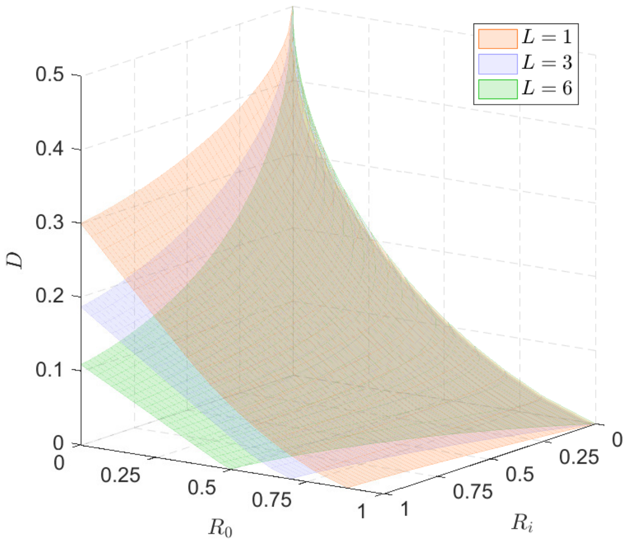
$$\frac{C(\gamma_0)}{r_0} \ge \begin{cases} g(D), & \text{for } 0 \le D \le D_c,\\ (D - D')g'(D_c), & \text{for } D_c < D \le D',\\ 0, & \text{for } D' < D,\end{cases} \quad (52)$$

with $\alpha_i = p_i * h^{-1}(1 - [C(\gamma_i)/r_i]^-)$ for calculating $g(D)$ and D'.

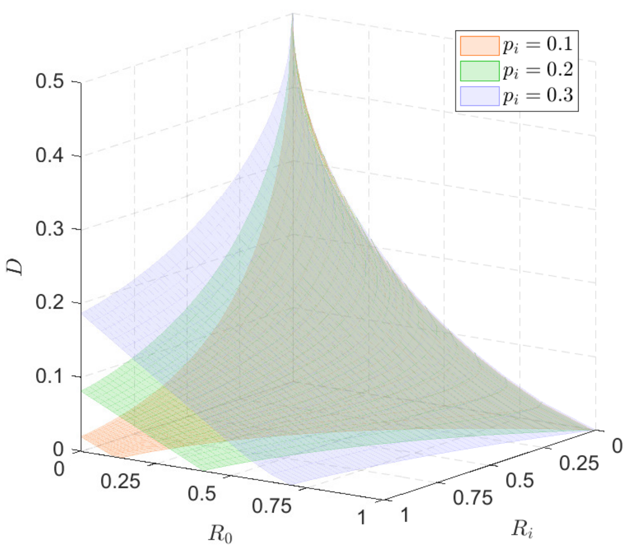
Remark: If a distortion requirement is given, we can evaluate whether the SNR of all links can satisfy the distortion requirement by (52). Conversely, if the SNR values of all links are given, we can utilize (52) to calculate the final distortion.

B. Numerical Results

The relationship between the link rates and the final distortion is illustrated in Fig. 5, where we set all R_i at the same value, i.e., homogeneous assisting links, so that the achievable rate-distortion region is able to be plotted within

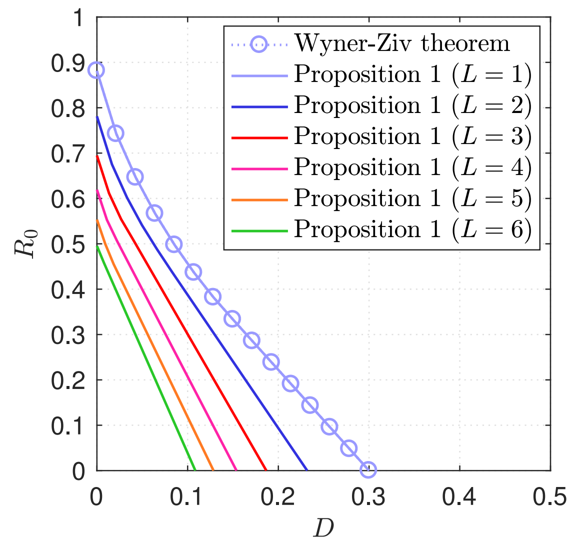


(a) $p_i = 0.3$.

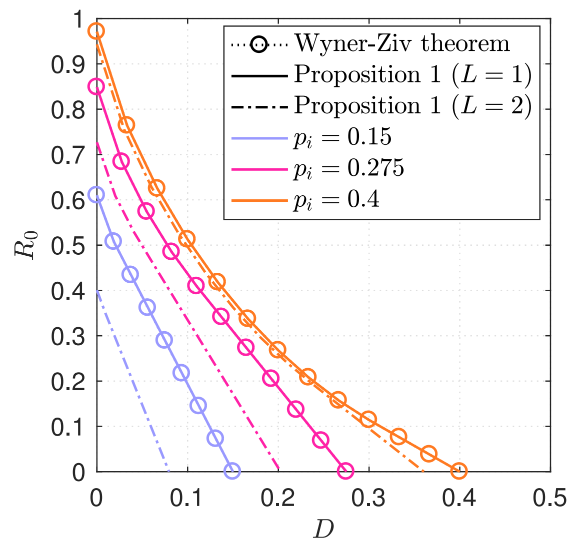


(b) $L = 3$.

Fig. 5. The achievable rate-distortion region for homogeneous assisting links.



(a) $p_i = 0.3$.



(b) Diverse p_i.

Fig. 6. The rate-distortion function for given $R_i = 1$.

three dimensions. From the whole view, we can see that the distortion of X drops from 0.5 as R_0 and R_i gradually increase from 0. Moreover, the distortion decreases faster for larger L and smaller p_i in Fig. 5(a) and Fig. 5(b), respectively. It is also remarkable that all surfaces of the rate-distortion function intersect at one same curve in the R_0-D coordinate plane, i.e., $R_i = 0$. Obviously, the system model is equivalent to independent lossy source coding if $R_i = 0$, and hence $R_0(D)$ reduces to the classical rate-distortion function, which is not affected by the number of assisting links and the correlations between sources. Another important phenomenon is that the distortion cannot be entirely eliminated to zero in the R_i-D coordinate plane. Therefore, the estimate $\hat{X}$ must be a lossy version of X when there is no information of X directly available for $p_i > 0$.

For given R_i, we can obtain the curves shown in Fig. 6 by projecting the surfaces of the rate-distortion function onto the R_0-D coordinate plane. Interestingly, the curves based on *Proposition 1* perfectly coincide with the curves of the Wyner-Ziv theorem for arbitrary p_i if there is only one assisting link without rate constraint. This phenomenon results from the fact that the theoretical model of the lossy source coding with helpers reduces to the Wyner-Ziv problem when $L = 1$ and $R_i = 1$. In addition, Fig. 6(a) demonstrates that the distortion can be reduced by introducing extra assisting links; however, the gap between L and $(L+1)$ becomes narrower along with the increment of assisting links. Consequently, it is harder to obtain more gains when the number of assisting links is already large enough. In Fig. 6(b), we can clearly observe that the curve shift to the left for small p_i, i.e., the distortion is smaller for more correlated sources. Meanwhile, the gap between L and $(L+1)$ is wider for the sources with high correlations, and hence it is more efficient to introduce extra assisting links for more correlated sources.

V. Performance Evaluation

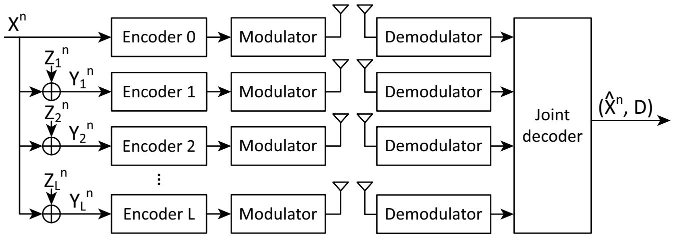
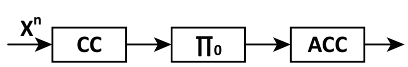
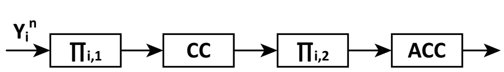


Fig. 7. An instance of cooperative communications systems in WSNs.

In this section, we start to evaluate the practical system performance for an instance of cooperative communications WSNs depicted in Fig. 7. There is one target sequence X^n and L assisting sequences Y_i^n corrupted by Z_i^n with $Z_i \sim \mathrm{Bern}(p_i)$. To begin with, encoder 0 and encoder i encode their own sequence, respectively, and send the codeword through additive white Gaussian noise (AWGN) channels after modulation. The objective of this simulation is to compare the practical performance with the theoretical bound. Therefore, in order to make the final distortion as small as possible, the fusion center starts to decode and produce estimate $\hat{X}^n$ after receiving and demodulating the signals in all the links. If the aim of a system is to satisfy a specified distortion requirement, the fusion center may decrease the latency and complexity by decoding with fewer assisting sequences. In other words, after receiving signals from some links, the fusion center can first evaluate the SNR of received signals and the crossover probabilities between X and Y_i by the error probability estimation algorithm proposed in [17]. Then, it calculates the final distortion with already received signals by (52). If the expected final distortion is not larger than the given distortion requirement, the fusion center starts decoding process; otherwise, it continues receiving the signals from the remaining links until the expected final distortion is small enough.



(a) The encoder of primary sensor.



(b) The encoder of i-th assisting sensor.

Fig. 8. The structure of encoders.

Since the distributed compress-bin scheme utilized in the theoretical proof requires extremely huge memory to store the codebook, we design a practical coding scheme for simulation. Fig. 8 illustrates the structure of encoders, which consists of two component codes, i.e., a convolutional code (CC) as the outer code and the output of an accumulator (ACC) [20] as the inner code. In order to exploit the principle of turbo code [21] in decoding, an interleaver Π_0 or $\Pi_{i,2}$ is deployed between CC and ACC in the primary link or the i-th assisting link, respectively. Moreover, an additional interleaver $\Pi_{i,1}$ is used to disperse noises into different bits before CC in the assisting link.

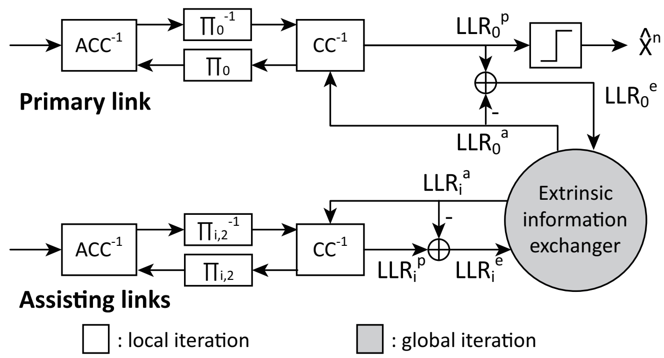


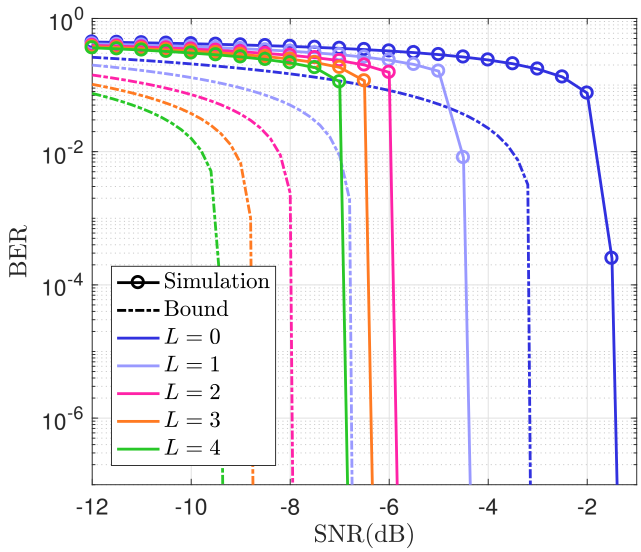
Fig. 9. The structure of the joint decoder.

As depicted in Fig. 9, the decoder of ACC (ACC^{-1}) decodes the inner code, and then the decoder of CC (CC^{-1}) decodes the outer code after deinterleaving in Π^{-1}. Next, the extrinsic information is interleaved and subsequently exchanged to ACC^{-1} as the *a priori* information in local iteration. In the global iteration, the *a posteriori* LLR (LLR^p) output from CC^{-1} is updated via an extrinsic information exchanger, which inputs the extrinsic LLR (LLR^e) and outputs the *a priori* LLR (LLR^a). The extrinsic information exchanger calculates LLR^e by the LLR updating function $\mu(\cdot)$ for correlated sources [22] based on the correlation model [23].

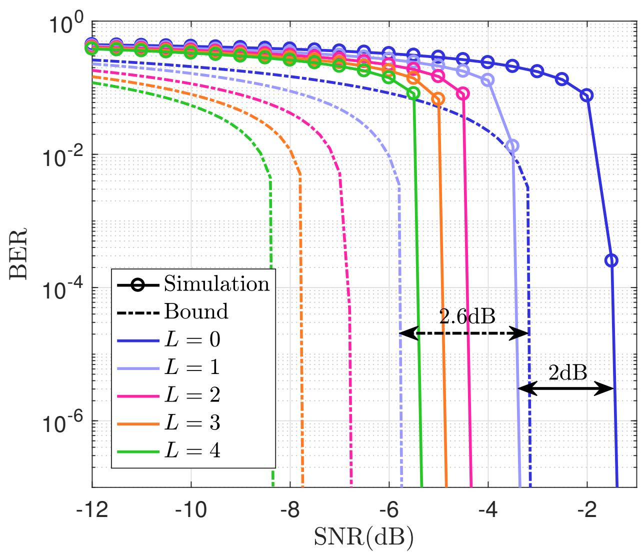
TABLE II
BASIC SETTINGS OF SIMULATION PARAMETERS

Parameter	Value
Number of Blocks	1000
Block length	10000 bits
Generator polynomial of CC	$G = ([3, 2]3)_8$
Rate of CC	1/2
Type of interleaver	random interleaver
Modulation method	BPSK
Maximum iteration time	30

With the basic parameter settings listed in Table II, the simulation results in Fig. 10 show the similar tendency as the curves of the theoretical bound. Clearly, the SNR threshold becomes lower as the number of assisting sensors increases; however, the turbo cliff shifts to the left less rapidly for the system with more assisting sensors. By the comparison between Fig. 10(a) and Fig. 10(b), we can find that the more independent the sources are, the higher SNR threshold is required. The performance gap between the theoretical and simulation results is due to the following two factors, i.e., the suboptimal channel coding scheme and incomplete utilization of joint typicality in the simulation. First, notice that there is an obvious gap between the theoretical and simulation results even for the case without any assisting sensor, because it is hard to achieve the Shannon limit by the relatively simple



(a) $p_i = 0.01$.



(b) $p_i = 0.1$.

Fig. 10. Simulation results, where SNR is set at the same value for all links.

channel coding scheme used in the simulation. Besides the loss of performance due to channel coding, another key factor for the gap between the theoretical and simulation results is the incomplete utilization of joint typicality in the simulation. For instance, as shown in Fig. 10(b), there is a 2.6 dB gain for the theoretical analysis between no assisting sensor and one assisting sensor; however, only 2 dB gain can be achieved in the simulation for the same condition. This observation implies that the joint typicality is not completely utilized for joint decoding in the simulation as the distributed compress-bin scheme used in the theoretical analysis.

VI. Conclusion

We have analyzed the performance of cooperative communications in WSNs for IoT, where the final distortion of the estimate is acceptable if the fusion center can still make right judgements and operations. To begin with, we start the theoretical analysis from a lossy source coding problem with helpers. After deriving an inner bound on the achievable rate-distortion region, we further calculate the rate-distortion function for binary sources. Subsequently, the results of multiterminal source coding is extended to joint source-channel coding based on Shannon's lossy source-channel separation theorem. The theoretical results perfectly match the Wyner-Ziv theorem, if there is only one assisting sensor and no rate limit on it. Finally, we present a distributed encoding and joint decoding scheme to evaluate the practical performance for an instance of cooperative communications systems in WSNs via a series of simulations. The comparison between the theoretical and simulation results inspires us that the system performance can be further improved if there is a better coding scheme which can more efficiently utilizes the joint typicality of the coded sequences. Moreover, both the theoretical and simulation results indicate that the additional assisting link provides even smaller gains as the number of the assisting links becomes large.

Appendix A
Proof of Lemma 1

First, consider

$$\begin{aligned}
&\mathrm{P}\{(\tilde{V}_{\mathcal{S}}^n, \tilde{v}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}, V_{\mathcal{S}^c})\} \\
&\quad = \sum_{\substack{\tilde{v}_{\mathcal{S}^c}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}^c}|\tilde{v}_{\mathcal{S}}^n), \\ \tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}})}} p(\tilde{v}_{\mathcal{S}^c}^n) \\
&\quad \le \mathrm{P}\{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}})\} \cdot |\mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}^c}|\tilde{v}_{\mathcal{S}}^n)| \\
&\qquad \cdot 2^{-n[H(V_{\mathcal{S}^c})-\epsilon H(V_{\mathcal{S}^c})]} \\
&\quad \le \mathrm{P}\{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}})\} \cdot 2^{n[H(V_{\mathcal{S}^c}|V_{\mathcal{S}})+\epsilon H(V_{\mathcal{S}^c}|V_{\mathcal{S}})]} \\
&\qquad \cdot 2^{-n[H(V_{\mathcal{S}^c})-\epsilon H(V_{\mathcal{S}^c})]} \\
&\quad = \mathrm{P}\{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}})\} \cdot 2^{-n[I(V_{\mathcal{S}};V_{\mathcal{S}^c})-\delta'(\epsilon)]} \\
&\quad = \mathrm{P}\{\tilde{v}_{S_1^{|\mathcal{S}|}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{|\mathcal{S}|}})\} \cdot 2^{-n[I(V_{\mathcal{S}};V_{\mathcal{S}^c})-\delta'(\epsilon)]}, \qquad (53)
\end{aligned}$$

where $\delta'(\epsilon) = \epsilon[H(V_{\mathcal{S}^c}) + H(V_{\mathcal{S}^c}|V_{\mathcal{S}})]$. To further calculate (53), consider

$$\begin{aligned}
&\mathrm{P}\{\tilde{v}_{S_1^j}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^j})\} \\
&\quad = \mathrm{P}\{(\tilde{v}_{S_1^{j-1}}^n, \tilde{v}_{S_j}^n) \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{j-1}}, V_{S_j})\} \\
&\quad = \sum_{\substack{\tilde{v}_{S_j}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_j}|\tilde{v}_{S_1^{j-1}}^n), \\ \tilde{v}_{S_1^{j-1}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{j-1}})}} p(\tilde{v}_{S_j}^n) \\
&\quad \le \mathrm{P}\{\tilde{v}_{S_1^{j-1}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{j-1}})\} \cdot |\mathcal{T}_\epsilon^{(n)}(V_{S_j}|\tilde{v}_{S_1^{j-1}}^n)| \\
&\qquad \cdot 2^{-n[H(V_{S_j})-\epsilon H(V_{S_j})]} \\
&\quad \le \mathrm{P}\{\tilde{v}_{S_1^{j-1}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{j-1}})\} \\
&\qquad \cdot 2^{n[H(V_{S_j}|V_{S_1^{j-1}})+\epsilon H(V_{S_j}|V_{S_1^{j-1}})]} \cdot 2^{-n[H(V_{S_j})-\epsilon H(V_{S_j})]} \\
&\quad = \mathrm{P}\{\tilde{v}_{S_1^{j-1}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^{j-1}})\} \\
&\qquad \cdot 2^{-n[I(V_{S_1^{j-1}};V_{S_j})-\delta_j'(\epsilon)]}, \qquad (54)
\end{aligned}$$

where $\delta'_j(\epsilon) = \epsilon[H(V_{S_j}) + H(V_{S_j}|V_{S_1^{j-1}})]$ and $j \in \{3, 4, \cdots, |\mathcal{S}|\}$. According to the joint typicality lemma, for $j = 2$, we have

$$\begin{aligned} &\mathrm{P}\{\tilde{v}^n_{S_1^2} \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^2})\} \\ &= \mathrm{P}\{(\tilde{v}^n_{S_1}, \tilde{v}^n_{S_2}) \in \mathcal{T}_\epsilon^{(n)}(V_{S_1}, V_{S_2})\} \\ &\le 2^{-n[I(V_{S_1};V_{S_2}) - \delta'_2(\epsilon)]}. \end{aligned} \tag{55}$$

By combining the results of (54) and (55), we have

$$\begin{aligned} &\mathrm{P}\{\tilde{v}^n_{S_1^j} \in \mathcal{T}_\epsilon^{(n)}(V_{S_1^j})\} \\ &\le \mathrm{pow}\left(2, -n\Big[\sum_{i=2}^{j}(I(V_{S_1^{i-1}}; V_{S_i}) - \delta'_i(\epsilon))\Big]\right). \end{aligned} \tag{56}$$

By substituting (56) into (53), we have

$$\begin{aligned} &\mathrm{P}\{(\tilde{V}^n_{\mathcal{S}}, \tilde{v}^n_{\mathcal{S}^c}) \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}, V_{\mathcal{S}^c})\} \\ &\le \mathrm{pow}\Bigg(2, -n\Big[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) \\ &\qquad - \delta(\epsilon)\Big]\Bigg), \end{aligned} \tag{57}$$

where $\delta(\epsilon) = \delta'(\epsilon) + \sum_{j=2}^{|\mathcal{S}|} \delta'_j(\epsilon)$. This completes the proof of *Lemma 1*.

Appendix B
Proof of Lemma 2

Define the events

$$\tilde{\mathcal{E}}_k = \{(V^n_{S_1}(k_{S_1}), \cdots, V^n_{S_{|\mathcal{S}|}}(k_{S_{|\mathcal{S}|}}), \tilde{V}^n_{\mathcal{S}^c}) \in \mathcal{T}_\epsilon^{(n)}\} \text{ for } k_i \in \mathcal{K}_i, i \in \mathcal{S}. \tag{58}$$

By the union of events bound, the probability of the event of interest can be bounded as

$$\begin{aligned} &\mathrm{P}\left(\bigcup_{k_i \in \mathcal{K}_i, i \in \mathcal{S}} \tilde{\mathcal{E}}_k\right) \\ &\le \sum_{k_i \in \mathcal{K}_i, i \in \mathcal{S}} \mathrm{P}(\tilde{\mathcal{E}}_k) \\ &= \prod_{i \in \mathcal{S}} 2^{nr_i} \cdot \mathrm{P}(\tilde{\mathcal{E}}_k) \\ &\le \prod_{i \in \mathcal{S}} 2^{nr_i} \cdot \mathrm{pow}\Bigg(2, -n\Big[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) \\ &\qquad - \delta(\epsilon)\Big]\Bigg) \end{aligned} \tag{59}$$

$$\begin{aligned} &= \mathrm{pow}\Bigg(2, n\sum_{i \in \mathcal{S}} r_i - n\Big[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) \\ &\qquad - \delta(\epsilon)\Big]\Bigg), \end{aligned} \tag{60}$$

where (59) follows according to *Lemma 1*. Notice that (60) tends to zero as $n \to \infty$ if

$$\sum_{i \in \mathcal{S}} r_i < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon). \tag{61}$$

This completes the proof of *Lemma 2*.

Acknowledgment

This work is funded in part by China Scholarship Council (CSC) and in part by JAIST Core-to-Core Program. This work has been also performed in part under JSPS Kakenhi (B)15H04007.

References

[1] A. Alaiad and L. Zhou, "Patients' adoption of WSN-based smart home healthcare systems: An integrated model of facilitators and barriers," *IEEE Transactions on Professional Communication*, vol. 60, no. 1, pp. 4–23, Mar. 2017.

[2] Q. Chi, H. Yan, C. Zhang, Z. Pang, and L. Da Xu, "A reconfigurable smart sensor interface for industrial WSN in IoT environment," *IEEE transactions on industrial informatics*, vol. 10, no. 2, pp. 1417–1425, May 2014.

[3] S. R. Nimbargi, S. Hadawale, and G. Ghodke, "Tsunami alert & detection system using IoT: A survey," in *International Conference on Big Data, IoT and Data Science*. Pune, India: IEEE, Dec. 2017, pp. 182–184.

[4] B. Siregar, A. B. A. Nasution, and F. Fahmi, "Integrated pollution monitoring system for smart city," in *International Conference on ICT For Smart Society (ICISS)*. Surabaya, Indonesia: IEEE, Jul. 2016, pp. 49–52.

[5] Z. Zou, Y. Bao, F. Deng, and H. Li, "An approach of reliable data transmission with random redundancy for wireless sensors in structural health monitoring," *IEEE Sensors Journal*, vol. 15, no. 2, pp. 809–818, Feb. 2015.

[6] S. Long and P. Xiang, "Lossless data compression for wireless sensor networks based on modified bit-level RLE," in *8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*. Shanghai, China: IEEE, Sep. 2012, pp. 1–4.

[7] V. Dedeoglu, S. Perreau, and A. Grant, "Distributed energy consumption minimization for lossless data gathering wireless sensor networks," in *Australian Communications Theory Workshop (AusCTW)*. Wellington, New Zealand: IEEE, Feb. 2012, pp. 145–149.

[8] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[9] ——, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, p. 1, Mar. 1959.

[10] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.

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

[12] R. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Transactions on Information Theory*, vol. 21, no. 6, pp. 629–637, Nov. 1975.

[13] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1978, pp. 171–231.

[14] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, New York, 1978.

[15] T. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Transactions on Information Theory*, vol. 26, no. 3, pp. 277–288, May 1980.

[16] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[17] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Estimation of observation error probability in wireless sensor networks," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1073–1076, Jun. 2013.

[18] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.

[19] X. He, X. Zhou, P. Komulainen, M. Juntti, and T. Matsumoto, "A lower bound analysis of hamming distortion for a binary CEO problem with joint source-channel coding," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 343–353, Jan. 2016.

[20] K. Anwar and T. Matsumoto, "Accumulator-assisted distributed turbo codes for relay systems exploiting source-relay correlation," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1114–1117, Jul. 2012.

[21] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[22] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Transactions on Communications*, vol. 95, no. 12, pp. 3654–3662, Dec. 2012.

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

Wensheng Lin (S'17) received the B.Eng. degree in communication engineering, and the M.Eng. degree in electronic and communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree at the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan. His research interests include network information theory, distributed source coding, helper structure design, and optimal encoding/decoding design.

Tad Matsumoto (S'84-SM'95-F'10) received the B.S., M.S., and Ph.D. degrees from Keio University, Yokohama, Japan, in 1978, 1980, and 1991, respectively, all in electrical engineering.

He joined Nippon Telegraph and Telephone Corporation (NTT) in April 1980. In July 1992, he transferred to NTT DoCoMo, Tokyo, Japan, where he researched Code-Division Multiple-Access techniques for Mobile Communication Systems. In April 1994, he transferred to NTT America, where he served as a Senior Technical Advisor of a joint project between NTT and NEXTEL Communications. In March 1996, he returned to NTT DoCoMo, where he served as the Head of the Radio Signal Processing Laboratory until August of 2001. He worked on adaptive signal processing, multiple-input multiple-output turbo signal detection, interference cancellation, and space-time coding techniques for broadband mobile communications. In March 2002, he moved to University of Oulu, Finland, where he served as a Professor with Centre for Wireless Communications. In 2006, he served as a Visiting Professor at Ilmenau University of Technology, Ilmenau, Germany, funded by the German MERCATOR Visiting Professorship Program. Since April 2007, he has been serving as a Professor with Japan Advanced Institute of Science and Technology (JAIST), Japan, while also keeping a cross appointment professorship at University of Oulu. He was appointed a Finland Distinguished Professor for a period from January 2008 to December 2012, funded by the Finnish National Technology Agency (Tekes), Helsinki, Finland, and Finnish Academy, under which he preserves the rights to participate in and apply to European and Finnish national projects. He was the recipient of the IEEE VTS Outstanding Service Award (2001), Nokia Foundation Visiting Fellow Scholarship Award (2002), IEEE Japan Council Award for Distinguished Service to the Society (2006), IEEE Vehicular Technology Society James R. Evans Avant Garde Award (2006), and Thuringen State Research Award for Advanced Applied Science (2006), 2007 Best Paper Award of Institute of Electrical, Communication, and Information Engineers of Japan (2008), Telecom System Technology Award by the Telecommunications Advancement Foundation (2009), IEEE Communications Letters Exemplifying Reviewer Award (2011), and Nikkei Wireless Japan Award (2012). He is a member of IEICE. He served as an IEEE Vehicular Technology Distinguished Lecturer during the term July 2011-June 2015.