Successive Wyner-Ziv Coding for the Binary CEO Problem Under Logarithmic Loss

Mahdi Nangir, Reza Asvadi, *Member, IEEE*, Jun Chen, *Senior Member, IEEE*, Mahmoud Ahmadian-Attari, and Tad Matsumoto, *Fellow, IEEE*

***Abstract*— The L-link binary Chief Executive Officer (CEO) problem under logarithmic loss is investigated in this paper. A quantization splitting technique is applied to convert the problem under consideration to a $(2L-1)$-step successive Wyner-Ziv (WZ) problem, for which a practical coding scheme is proposed. In the proposed scheme, Low-Density Generator-Matrix (LDGM) codes are used for binary quantization while Low-Density Parity-Check (LDPC) codes are used for syndrome generation; the decoder performs successive decoding based on the received syndromes and produces a soft reconstruction of the remote source. The simulation results indicate that the rate-distortion performance of the proposed scheme can approach the theoretical inner bound based on binary-symmetric test-channel models.**

***Index Terms*— Binary CEO problem, binary quantization, successive decoding, syndrome decoding, Wyner-Ziv problem, quantization splitting, logarithmic loss.**

I. Introduction

MULTITERMINAL source coding is an important subject of network information theory. Research on this subject has yielded insights and techniques that are useful for a wide range of applications, including, among other things, cooperative communications [2], distributed storage [3], and sensor networks [4]. A particular formulation of multiterminal source coding, known as the Chief Executive Officer (CEO) problem, has received significant attention [5]. In this problem, there are L encoders (also called agents), which observe independently corrupted versions of a source; these encoders compress their respective observations and forward the compressed data separately to a central decoder (also called CEO), which then produces a (lossy) reconstruction of the target source.

The quadratic Gaussian setting of the CEO problem has been studied extensively, for which the rate-distortion region is characterized completely in the scalar case and partially in the vector case. Extending these results beyond the quadratic Gaussian setting turns out to be highly non-trivial; there are some results in [6]–[8]. Indeed, even for many seemingly simple sources and distortion measures, the understanding of the relevant information-theoretic limits is rather limited. A remarkable exception is a somewhat under-appreciated distortion measure called logarithmic loss (log-loss). As shown by Courtade and Weissman [9], the rate-distortion region of the CEO problem under log-loss admits a single-letter characterization for arbitrary finite-alphabet sources and noisy observations. Different from the conventional distortion measures which are typically imposed on "hard" reconstructions defined over the given source alphabet, log-loss is tailored to "soft" reconstructions in the form of probability distributions. Specifically, in the context of the CEO problem, the most favorable "soft" reconstruction is essentially the *a posteriori* distribution of the source given the compressed data received from the encoders (which is a sufficient statistic); it is more informative than its "hard" counterparts and more suitable for many downstream statistical inference tasks.

Recent years have seen significant interests in a new paradigm of wireless communications called Cloud-Radio Access Network (C-RAN). It has been recognized that the information-theoretic and coding-theoretic aspect of C-RAN is closely related to that of the CEO problem under log-loss [10]. This intriguing connection greatly enriches the implication of the latter problem and provides further motivations for the relevant research.

A main contribution of the present paper is a practical coding scheme for the CEO problem under log-loss. We adopt a hierarchical approach by decomposing the CEO problem into a set of simpler problems upon which the existing coding techniques can be directly brought to bear and then

Manuscript received December 30, 2018; revised May 26, 2019, July 11, 2019 and August 2, 2019; accepted August 21, 2019. Date of publication August 27, 2019; date of current version November 19, 2019. This work was supported in part by the project 7th Framework Program (FP7) Links-on-the-fly Technology for Robust, Efficient and Smart Communication in Unpredictable Environments (RESCUE), and in part by JAIST Core-to-Core program. This paper was presented in part at the 29th Biennial Symposium on Communications (BSC 2018), Toronto, Canada [1]. The associate editor coordinating the review of this article and approving it for publication was R. Thobaben. *(Corresponding author: Reza Asvadi.)*

M. Nangir was with the Faculty of Electrical Engineering, K. N. Toosi University of Technology, Tehran 1969764499, Iran. He is now with the Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 5166616471, Iran (e-mail: nangir@tabrizu.ac.ir).

R. Asvadi is with the Faculty of Electrical Engineering, Shahid Beheshti University, Tehran 1983969411, Iran (e-mail: r_asvadi@sbu.ac.ir).

J. Chen is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada (e-mail: junchen@mail.ece.mcmaster.ca).

M. Ahmadian-Attari is with the Faculty of Electrical Engineering, K. N. Toosi University of Technology, Tehran 1969764499, Iran (e-mail: mahmoud@eetd.kntu.ac.ir).

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

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

Digital Object Identifier 10.1109/TCOMM.2019.2937932

TABLE I
LIST OF SOME SYMBOLS USED IN THIS PAPER

Symbol	Description
q_j	Appearance probability of the binary representation of j in the links' output
Q_j	Appearance probability of the binary representation of j in the links' output for a specific $x \in \mathbb{B}$
p_l	Bernoulli noise parameter in the l-th link
Y_l^n	Binary noisy observation in the l-th link
d_l	Crossover probability of the binary symmetric test channel in the l-th link
P_l	Hamming distance between remote source and quantized sequence in the l-th link
$\phi_i(\cdot)$	Mapping from $\mathbb{B}^{K_i}$ to $\mathbb{B}$
$\mu_{\max}$	Minimum value of μ for the case of $R_{\text{th}} = 0$ and $D_{\text{th}} = 1$
k_i and k_i'	Number of syndrome bits in compound LDGM-LDPC codes
m_l	Number of variable nodes in LDGM code $\mathcal{C}_{\boldsymbol{W}_l}$ of the l-th link
M_i	Number of variable nodes in LDGM code $\mathcal{C}_i'$ of the i-th link
N_l^n	Observation noise in the l-th link
W_l	Output sequence of splitter in the l-th link
U_l	Quantized sequence in the l-th link
X^n	Remote binary symmetric source (n-tuple)
μ	Slope of the tangent line to sum rate-distortion bound curve
$\hat{X}^n$	Soft reconstruction of the remote source
δ_l	Splitting parameter in the l-th link
$R_{\text{th}}(D_{\text{th}})$	Theoretical bound of Sum-rate (Distortion)
σ	Well-ordered permutation

combining these small pieces to find the solution to the original problem. Two most basic problems in information theory are point-to-point channel coding and (lossy) source coding (also known as quantization). It is well known that the fundamental limits of these two problems can be approached using graph-based codes (e.g., Low-Density Parity-Check (LDPC) codes for channel coding [11] and Low-Density Generator-Matrix (LDGM) codes for (lossy) source coding [12]) in conjunction with iterative message-passing algorithms (e.g., the Sum-Product (SP) algorithm for channel decoding [11] and the Bias-Propagation (BiP) algorithm for (lossy) source encoding [13], [14]). These basic coding components can serve as the building blocks of more sophisticated schemes for the problems at the second level of the hierarchy. Notable examples include the Gelfand-Pinsker problem and the Wyner-Ziv problem, which are solved via proper combination of source codes and channel codes [15], [16]. With these solutions in hand, one can then tackle the problems at the third level or even higher. From this perspective, our proposed scheme for the CEO problem can be interpreted as successive implementation of Wyner-Ziv coding.

The conversion of the CEO problem to the Wyner-Ziv problem is realized using quantization splitting. The idea of quantization splitting is by no means new. Indeed, it has been applied to the multiterminal source coding problem [17] and multiple description problem [18] among others [19], particularly in the quadratic Gaussian setting. However, to the best of our knowledge, the application of quantization splitting is mainly restricted in the theoretical domain as a conceptual apparatus, and its practical implementation has not been addressed in the literature, at least for the problem under consideration (namely, the CEO problem under log-loss). In this work we mainly focus on the setting where the source is binary-symmetric and is corrupted by independent Bernoulli noises. It is worth emphasizing that this simple setting captures the essential features of the CEO problem and the methodology underlying our proposed scheme is in fact broadly applicable.

The organization of this paper is as follows. The problem definition and the concept of quantization splitting are presented in Section II. The proposed scheme is described in Section III. Sections IV and V contain associated analytical and numerical results, respectively. We conclude the paper in Section VI.

II. THE CEO PROBLEM AND QUANTIZATION SPLITTING

A. Notations

Throughout this paper, the logarithm is to the base 2. Random variables and their realizations are shown by capital letters and lowercase letters in italics, respectively. Sets and alphabet set of random variables are depicted by calligraphic letters. Furthermore, matrices are shown by bold-faced letters. The binary entropy function is $h_b(x) = -x\log x - (1-x)\log(1-x)$, $\mathbb{B} \triangleq \{0,1\}$, and $p * d = p(1-d) + (1-p)d$ shows the binary convolution of p and d. The list of some symbols used in the paper is represented in Table I.

B. System Model

Let $X^n = (X_1, \cdots, X_n)$ be an independent and identically distributed (i.i.d.) remote source. L noisy observations of X^n are available in L links that are mutually independent without any communication among them. These noisy observations, Y_l^n for $l \in \mathcal{I}_L \triangleq \{1, \cdots, L\}$, are generated by X^n through independent memoryless channels. The block diagram of an L-link CEO problem is depicted in Fig. 1. In each link, an encoder maps its noisy observation to a codeword C_l by using a function f_l, as follows:

$$C_l = f_l(Y_l^n), \quad \text{where } Y_l^n \in \mathcal{Y}_l^n \text{ and } C_l \in \mathcal{C}_l, \text{ for } l \in \mathcal{I}_L. \tag{1}$$

The codewords C_l, for $l \in \mathcal{I}_L$, are sent to a joint CEO decoder via noiseless channels. The CEO decoder produces a soft reconstruction $\hat{X}^n = (\hat{X}_1, \cdots, \hat{X}_n)$ of the original remote

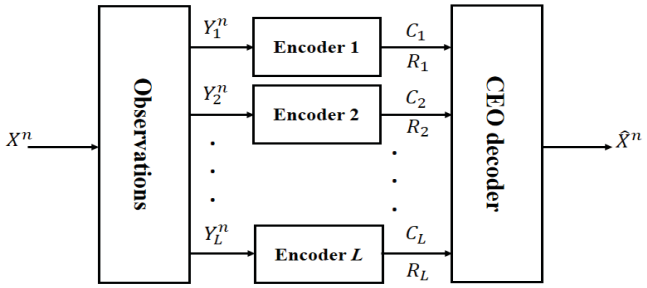


Fig. 1. Configuration of an L-link CEO problem.

source X^n by using a function g, as follows:

$$\hat{X}^n = g(C_1, \cdots, C_L), \quad \text{where } (C_1, \cdots, C_L) \in \mathcal{C}_1 \cdots \times \mathcal{C}_L. \tag{2}$$

Definition 1: The log-loss induced by a symbol $x \in \mathcal{X}$ and a probability distribution $\hat{x}$ on $\mathcal{X}$ is defined as

$$d(x, \hat{x}) = \log\big(\frac{1}{\hat{x}(x)}\big). \tag{3}$$

More generally, for a sequence of symbols $x^n = (x_1, \cdots, x_n)$ and a sequence of distributions $\hat{x}^n = (\hat{x}_1, \cdots, \hat{x}_n)$, let

$$d(x^n, \hat{x}^n) = \frac{1}{n}\sum_{t=1}^{n} \log\big(\frac{1}{\hat{x}_t(x_t)}\big). \tag{4}$$

Definition 2: A rate-distortion vector $(R_1, \cdots, R_L, D)$ is called strict-sense achievable under log-loss, if for all sufficiently large n, there exist functions f_1,f_2,...,f_L, and g respectively according to (1) and (2) such that

$$\begin{aligned} R_l &\geq \frac{1}{n}\log\big|\mathcal{C}_l\big|, \quad \text{for } l \in \mathcal{I}_L; \\ D &\geq \mathbb{E}(d(X^n, \hat{X}^n)), \end{aligned} \tag{5}$$

where $\mathbb{E}(\cdot)$ denotes expectation function. The closure of the set of all strict-sense achievable vectors $(R_1, \cdots, R_L, D)$ is called the rate-distortion region of the CEO problem under log-loss and is denoted by $\overline{\mathcal{RD}}^{\star}_{\text{CEO}}$.

Definition 3 ([9, Definition 7]): Let $\mathcal{RD}^{i}_{\text{CEO}}$ be the set of all $(R_1, \cdots, R_L, D)$ satisfying

$$\sum_{l\in\mathcal{A}} R_l \geq I(Y_{\mathcal{A}}; U_{\mathcal{A}}|U_{\mathcal{A}^c}, Q), \quad \emptyset \subset \mathcal{A} \subseteq \mathcal{I}_L, \tag{6a}$$

$$D \geq H(X|U_{\mathcal{I}_L}, Q), \tag{6b}$$

for some joint distribution

$$p_Q(q)p_X(x)\prod_{l=1}^{L} p_{Y_l|X}(y_l|x)p_{U_l|Y_l,Q}(u_l|y_l, q), \tag{7}$$

where in (6a), $Y_{\mathcal{A}} = \{Y_l : l \in \mathcal{A}\}$ and $\mathcal{A}^c = \mathcal{I}_L \backslash \mathcal{A}$.

Definition 4 ([9, Definition 8]): Let $\mathcal{RD}^{o}_{\text{CEO}}$ be the set of all $(R_1, \cdots, R_L, D)$ satisfying

$$\sum_{l\in\mathcal{A}} R_l \geq [\sum_{l\in\mathcal{A}} I(Y_l; U_l|X, Q) + H(X|U_{\mathcal{A}^c}, Q) - D]^+, \quad \emptyset \subset \mathcal{A} \subseteq \mathcal{I}_L, \tag{8}$$

and (6b), for some joint distribution (7), where $[x]^+ = \max\{0, x\}$ and $U_{\mathcal{A}} \leftrightarrow Y_{\mathcal{A}} \leftrightarrow X \leftrightarrow Y_{\mathcal{A}^c} \leftrightarrow U_{\mathcal{A}^c}$ forms a Markov chain for any $\mathcal{A} \subseteq \mathcal{I}_L$.

It is shown in [9] that

$$\overline{\mathcal{RD}}^{\star}_{\text{CEO}} = \mathcal{RD}^{i}_{\text{CEO}} = \mathcal{RD}^{o}_{\text{CEO}}; \tag{9}$$

moreover, there is no loss of generality in imposing the cardinality bounds $|\mathcal{U}_l| \leq |\mathcal{Y}_l|, l \in \mathcal{I}_L$ and $|\mathcal{Q}| \leq L + 2$ on the alphabet sizes of auxiliary random variables U_l and timesharing variable Q, respectively.

Given test channels $p_{U_l|Y_l}$, $l \in \mathcal{I}_L$, we define $\mathcal{RD}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ as the set of all $(R_1, \cdots, R_L, D)$ satisfying

$$\sum_{l\in\mathcal{A}} R_l \geq I(Y_{\mathcal{A}}; U_{\mathcal{A}}|U_{\mathcal{A}^c}), \quad \emptyset \subset \mathcal{A} \subseteq \mathcal{I}_L, \tag{10}$$

$$D \geq H(X|U_{\mathcal{I}_L}), \tag{11}$$

where X, $Y_{\mathcal{I}_L}$, and $U_{\mathcal{I}_L}$ are jointly distributed according to $p_X(x)\prod_{l=1}^{L} p_{Y_l|X}(y_l|x)p_{U_l|Y_l}(u_l|y_l)$.

Note that (10) and (11) respectively correspond to (6a) and (6b) with timesharing variable Q set to be a constant. Therefore, $\mathcal{RD}^{i}_{\text{CEO}}$ (as well as $\mathcal{RD}^{o}_{\text{CEO}}$ and $\overline{\mathcal{RD}}^{\star}_{\text{CEO}}$ in light of (9)) can be expressed as the convex hull of the union of $\mathcal{RD}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ over all $(p_{U_l|Y_l}, l \in \mathcal{I}_L)$. Moreover, we define $\mathcal{R}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ to be the set of all $(R_1, \cdots, R_L)$ satisfying (10) and define its dominant face, denoted by $\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$, to be the set of $(R_1, \cdots, R_L) \in \mathcal{R}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ satisfying $\sum_{l=1}^{L} R_l = I(Y_{\mathcal{I}_L}; U_{\mathcal{I}_L})$. Due to the contra-polymatroid structure of $\mathcal{R}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ [17], [19], $\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ is non-empty and every $(R_1, \cdots, R_L, D)$ in $\mathcal{RD}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ is dominated, in a component-wise manner, by $(R'_1, \cdots, R'_L, H(X|U_{\mathcal{I}_L}))$ for some $(R'_1, \cdots, R'_L) \in \mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$.

C. Quantization Splitting

$\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ has $L!$ corner points. Specifically, each permutation π on $\mathcal{I}_L$ is associated with a corner point $(R_1(\pi), \cdots, R_L(\pi))$ of $\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ as follows:

$$\begin{aligned} R_{\pi(l)}(\pi) &= I(Y_{\pi(l)}; U_{\pi(l)}|U_{\pi(l+1)}, \cdots, U_{\pi(L)}), \quad l \in \mathcal{I}_{L-1}, \\ R_{\pi(L)}(\pi) &= I(Y_{\pi(L)}; U_{\pi(L)}). \end{aligned}$$

These corner points can be achieved via successive Wyner-Ziv coding with decoding order $U_{\pi(L)} \to U_{\pi(L-1)} \to \cdots \to U_{\pi(1)}$ (an implementation of this scheme for the case $L = 2$ can be found in [20]).

To achieve non-corner points of $\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$, we employ the quantization splitting technique introduced in [17], which is a generalization of the source splitting technique [21] and a counterpart of the rate splitting technique in channel coding [22], [23]. Roughly speaking, the basic idea underlying the quantization splitting technique is that each non-corner point in the L-dimensional space can be projected to a corner point in the $(2L-1)$-dimensional space. Specifically, it is known [17, Theorem 2.1] that, for any rate tuple $(R_1, \cdots, R_L) \in \mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$, there exist

random variables W_l, $l \in \mathcal{I}_L$, and a well-ordered permutation σ[1] on the set $\{W_1, \cdots, W_L, U_1, \cdots, U_L\}$ such that

$$R_l = I(Y_l; W_l | \{W_l\}_\sigma^-) + I(Y_l; U_l | \{U_l\}_\sigma^-), \quad l \in \mathcal{I}_L, \quad (12)$$

where $\{W_l\}_\sigma^-$ and $\{U_l\}_\sigma^-$ represent the set of random variables that respectively appear before W_l and U_l in the well-ordered permutation σ; moreover, W_l is a physically degraded version U_l, $l \in \mathcal{I}_L$, and at least one W_l is independent of U_l (and thus can be eliminated).

It is instructive to view U_l as a fine-description of Y_l and view W_l as a coarse-description split from U_l, $l \in \mathcal{I}_L$. Eq. (12) suggests that the given rate tuple $(R_1, \cdots, R_L)$ can be achieved via successive Wyner-Ziv coding with decoding order specified by σ. It should be emphasized that the successive Wyner-Ziv coding scheme for non-corner points is in general more complicated than that for corner points. First of all, the scheme for non-corner points involves more encoding and decoding steps. Secondly and more importantly, to realize the splitting effect, one needs to generate a coarse-description codebook and then, for each of its codewords, generate a fine-description codebook; as a consequence, the number of fine-description codebooks grows exponentially with the codeword length, causing a serious problem in practice. In this work we circumvent this problem by using a codebook construction technique inspired by the functional representation lemma [24], [25]. Successive refinement coding scheme is also a multi-terminal encoding problem for, basically, downlink, where terminals are classified into several groups, each having different distortion requirements. The remote source is encoded such that the description for the groups having higher distortion requirement can help recover another groups having lower distortion requirement. Alternatively, our proposed coding scheme successively decodes binary observations and then softly reconstructs the remote source with a single value of distortion under the log-loss criterion.

III. Description of the Proposed Scheme

Consider an L-link binary CEO problem, where a remote Binary-Symmetric Source (BSS) is corrupted by independent Bernoulli noises with parameters p_1, p_2, ..., and p_L, i.e.,

$$X \sim \text{Ber}(\frac{1}{2}), \quad Y_l = X \oplus N_l, \ N_l \sim \text{Ber}(p_l), \ l \in \mathcal{I}_L. \quad (13)$$

We make the following two assumptions.

1) A binary-symmetric test channel model is adopted for each encoder. More specifically, it is assumed that $p_{U_l|Y_l}$ is a Binary-Symmetric Channel (BSC) with crossover probability d_l, $l \in \mathcal{I}_L$. Hence, we can write $U_l = Y_l \oplus Z_l$, $l \in \mathcal{I}_L$, where $Z_l \sim \text{Ber}(d_l)$, $l \in \mathcal{I}_L$, are mutually independent and are independent of $(X, Y_{\mathcal{I}_L})$ as well. This assumption is justified by the numerical results in [20].
2) A BSC model is adopted for each splitter. More specifically, it is assumed that $p_{W_l|U_l}$ is a BSC with crossover probability δ_l, $l \in \mathcal{I}_L$. Hence, we can write $W_l = U_l \oplus V_l$, $l \in \mathcal{I}_L$, where $V_l \sim \text{Ber}(\delta_l)$, $l \in \mathcal{I}_L$, are mutually independent and are independent of $(X, Y_{\mathcal{I}_L}, U_{\mathcal{I}_L})$ as well. According to [23, Definition 2], this assumption incurs no loss of generality.

Since the coding schemes associated with different well-ordered permutations are conceptually similar, for ease of exposition, we focus on a specific permutation $\sigma = (W_1, \cdots, W_{L-1}, U_L, U_{L-1}, \cdots, U_1)$ (we eliminate W_L by setting $\delta_L = \frac{1}{2}$). Each conditional mutual information in (12) can be written as the difference of two terms, one associated with quantization and the other with binning. As an example, consider the second term of R_1, i.e., $I(Y_1; U_1 | W_1, \cdots, W_{L-1}, U_2, \cdots, U_L)$. We have

$$\begin{aligned} &I(Y_1; U_1 | W_1, \cdots, W_{L-1}, U_2, \cdots, U_L) \\ &\quad = I(Y_1; U_1 | W_1, U_2, \cdots, U_L) \quad (14) \\ &\quad = I(U_2, \cdots, U_L, Y_1; U_1 | W_1) - I(U_2, \cdots, U_L; U_1 | W_1) \\ &\quad = I(Y_1; U_1 | W_1) - I(U_2, \cdots, U_L; U_1 | W_1), \quad (15) \end{aligned}$$

where (14) is due to the degradeness of W_l with respect to U_l, $l \in [2 : L-1]$, where $[j : k] \triangleq \{j, \cdots, k\}$, and (15) is because of the fact that (U_1, W_1) and $(U_2, \cdots, U_L)$ are conditionally independent given Y_1. The term $I(Y_1; U_1 | W_1)$ specifies the quantization rate needed to generate the fine-description U_1 given the coarse description W_1 while the term $I(U_2, \cdots, U_L; U_1 | W_1)$ specifies the amount of rate reduction achievable through binning.

We use a binary quantizer to map the outputs of a BSS to the codewords of an LDGM code with the minimum Hamming distance. These quantizers are utilized in the encoders of our proposed coding scheme. Practically, binary quantization can be realized by using some iterative message passing algorithms such as the BiP algorithm [13] or the survey-propagation algorithm [12]. Presence of side information can further reduce the compression rate required for a prescribed distortion constraint. Actually, this lossless source coding scenario can be practically realized by a binning operation based on channel coding schemes [4]. In our proposed coding scheme, binning is implemented by using LDPC codes with the syndrome generation scheme. This binning scheme is also used for the asymmetric Slepian-Wolf coding problem. In practice, the SP algorithm can be used to iteratively decode the LDPC coset code specified by the given syndrome.

A. The Proposed Coding Scheme: An Information-Theoretic Description

To elucidate the overall structure of the proposed scheme, we first give a short description using the information-theoretic terminology. First, let $W_L \triangleq U_L$. In the following description, all the ϵ quantities are small positive real numbers.

Codebook Generation:

1) For $l \in \mathcal{I}_L$, a codebook $\mathcal{C}_{W_l}$ of rate $I(Y_l; W_l) + \epsilon_{l,1}$ should be constructed with each codeword generated independently according to $\prod_{t=1}^n p_{W_l}(\cdot)$.
2) For $i \in \mathcal{I}_{L-1}$ and each codeword $w_i^n \in \mathcal{C}_{W_i}$, a codebook $\mathcal{C}_{U_i}(w_i^n)$ of rate $I(Y_i; U_i | W_i) + \epsilon_{i,2}$ is required with

[1] A well-ordered permutation is an arbitrary ordering of the set $\{W_1, \cdots, W_L, U_1, \cdots, U_L\}$ with W_l appearing before U_l for all $l \in \mathcal{I}_L$.

each codeword generated independently according to $\prod_{t=1}^{n} p_{U_i|W_i}(\cdot|w_{i,t})$.

3) For $i \in [2:L]$, the $\mathcal{C}_{W_i}$ should be partitioned into $2^{n[I(Y_i;W_i|W_1,\cdots,W_{i-1})+\epsilon_{i,3}]}$ bins, where each bin contains $2^{n[I(W_1,\cdots,W_{i-1};W_i)-\epsilon_{i,3}]}$ codewords.
4) For $i \in \mathcal{I}_{L-1}$ and each codeword $w_i^n \in \mathcal{C}_{W_i}$, $\mathcal{C}_{U_i}(w_i^n)$ should be partitioned into $2^{n[I(Y_i;U_i|W_1,\cdots,W_i,U_{i+1},\cdots,U_L)+\epsilon_{i,4}]}$ bins, where each bin contains $2^{n[I(W_1,\cdots,W_{i-1},U_{i+1},\cdots,U_L;U_i|W_i)-\epsilon_{i,4}]}$ codewords.

Encoding:

1) For $l \in \mathcal{I}_L$ and a given y_l^n, the l-th encoder finds a codeword $w_l^n \in \mathcal{C}_{W_l}$ that is jointly typical with y_l^n. Note that the Hamming distance between w_l^n and y_l^n is approximately $n(d_l * \delta_l)$.
2) For $i \in \mathcal{I}_{L-1}$, the i-th encoder finds a codeword $u_i^n \in \mathcal{C}_{U_i}(w_i^n)$ that is jointly typical with (y_i^n, w_i^n). Note that the Hamming distance between u_i^n and y_i^n is approximately nd_i while the Hamming distance between u_i^n and w_i^n is approximately $n\delta_i$.
3) For $l \in \mathcal{I}_L$, the l-th encoder sends the index $b(w_l^n)$ of the bin that contains w_l^n (for $l = 1$, it only sends the index $i(w_1^n)$ of w_1^n, and for $l = L$ nothing is sent), and the index $b(u_l^n)$ of the bin that contains u_l^n to the decoder.

Decoding:

1) The decoder first decodes w_1^n based on $i(w_1^n)$.
2) For $i \in [2:L]$, it decodes w_i^n by searching in the bin with index $b(w_i^n)$ for the unique codeword that is jointly typical with $(w_1^n, w_2^n, \cdots, w_{i-1}^n)$.
3) For $j \in [L-1:1]$, it decodes u_j^n by searching in the bin with index $b(u_j^n)$ for the unique codeword that is jointly typical with $(w_1^n, \cdots, w_j^n, u_{j+1}^n, \cdots, u_L^n)$.
4) Finally, it uses $(\hat{u}_1^n, \cdots, \hat{u}_L^n)$ to produce a soft reconstruction of x^n by the following rule:

$$\hat{x}_t = p_{X|U_{\mathcal{I}_L}}(\cdot|\hat{u}_{1,t}, \cdots, \hat{u}_{L,t}), \quad t \in \mathcal{I}_n. \tag{16}$$

The conditional probability function $p_{X|U_{\mathcal{I}_L}}(\cdot|\hat{u}_{1,t}, \cdots, \hat{u}_{L,t})$ depends on the binary values of $\hat{u}_{l,t}$, for $l \in \mathcal{I}_L$. This function can be determined based on the joint distribution diagram of the CEO problem. As an example for $L = 3$, we calculate $p_{X|U_1,U_2,U_3}(x_j|\hat{u}_{1,j}, \hat{u}_{2,j}, \hat{u}_{3,j})$ for $(x_j = 0, \hat{u}_{1,j} = 1, \hat{u}_{2,j} = 0, \hat{u}_{3,j} = 1)$. The other 15 cases can be calculated similarly.

$$\begin{aligned}\hat{x}_j &= p_{X|U_1,U_2,U_3}(0|1,0,1) = \frac{p_{X,U_1,U_2,U_3}(0,1,0,1)}{p_{U_1,U_2,U_3}(1,0,1)} \\ &= \frac{p_{U_1,U_2,U_3|X}(1,0,1|0) \times 0.5}{p_{U_1,U_2,U_3|X}(1,0,1|0)\times 0.5 + p_{U_1,U_2,U_3|X}(1,0,1|1)\times 0.5} \\ &= \frac{\tau_1}{\tau_1+\tau_2},\end{aligned} \tag{17}$$

where $\tau_1 = (p_1 * d_1)(1 - p_2 * d_2)(p_3 * d_3)$ and $\tau_2 = (1 - p_1 * d_1)(p_2 * d_2)(1 - p_3 * d_3)$.

B. The Proposed Coding Scheme: A Coding-Theoretic Description

Now we translate the above information-theoretic description of the proposed scheme to a coding-theoretic description. Along the way, we address certain practical issues encountered in codebook generation using a construction technique inspired by the functional representation lemma. For notational simplicity, the description is given for the case $L = 3$; the extension to the general case is straightforward.

Codebook Generation:

1) For $l \in \mathcal{I}_3$, an LDGM codebook $\mathcal{C}_{W_l}$ should be generated with the rate of $I(Y_l; W_l) + \epsilon_{l,1} = 1 - h_b(d_l * \delta_l) + \epsilon_{l,1}$.[2]
2) For $i \in \mathcal{I}_2$ and each codeword w_i^n, a codebook $\mathcal{C}_{U_i}(w_i^n)$ is constructed as follows[3]:
Firstly, an LDGM code $\mathcal{C}_i'$ should be considered with $2^{n[I(Y_i;U_i|W_i)+\epsilon_{i,2}]}$ codewords with each of length nK_i, where K_i is a fixed integer. Let $\phi_i(\cdot)$ be a mapping[4] from $\mathbb{B}^{K_i} \to \mathbb{B}$ such that

$$|S_0| \approx 2^{K_i}(1-\delta_i) \quad \text{and } |S_1| \approx 2^{K_i}\delta_i, \tag{18}$$

where $S_b \triangleq \{s^{K_i} \in \mathbb{B}^{K_i} : \ \phi_i(s^{K_i}) = b\}$, $b \in \mathbb{B}$. Note that the approximation in (18) can be made arbitrarily precise when $K_i \to \infty$. For each codeword $c^{nK_i} \triangleq (c_1, \cdots, c_{nK_i}) \in \mathcal{C}_i'$, the c^{nK_i} is mapped to a codeword of length n by using $\phi_i(\cdot)$ as below:

$$\Big(\phi_i\big(c_1, \cdots, c_{K_i}\big), \cdots, \phi_i\big(c_{(n-1)K_i+1}, \cdots, c_{nK_i}\big)\Big). \tag{19}$$

By doing this for all codewords in $\mathcal{C}_i'$, a new codebook $\phi_i(\mathcal{C}_i')$ is obtained with $2^{n[I(Y_i;U_i|W_i)+\epsilon_{i,2}]}$ codewords, each of length n. Hence, the codebook $\mathcal{C}_{U_i}(w_i^n)$ can be defined as $w_i^n \oplus \phi_i(\mathcal{C}_i')$, which is a codebook obtained by adding w_i^n to each codeword in $\phi_i(\mathcal{C}_i')$. Now consider the backward channels $Y_i = U_i \oplus Z_i'$ and $U_i = W_i \oplus V_i'$,[5] where $Z_i' \sim \text{Ber}(d_i)$, $V_i' \sim \text{Ber}(\delta_i)$, and W_i are mutually independent. It can be verified that

$$I(Y_i; U_i|W_i) = I(V_i' \oplus Z_i'; V_i') = h_b(\delta_i * d_i) - h_b(d_i). \tag{20}$$

3) For $i = 2, 3$, to partition $\mathcal{C}_{W_i}$ into $2^{n[I(Y_i;W_i|W_1,\cdots,W_{i-1})+\epsilon_{i,3}]}$ bins with each bin containing $2^{n[I(W_1,\cdots,W_{i-1};W_i)-\epsilon_{i,3}]}$ codewords, an LDPC code of rate $I(W_1, \cdots, W_{i-1}; W_i) - \epsilon_{i,3}$ is used with parity-check matrix $\mathbf{H}_i = (\tilde{\mathbf{H}}_i, \hat{\mathbf{H}}_i)$, where $\tilde{\mathbf{H}}_i$ is the parity-check matrix of $\mathcal{C}_{W_i}$. It can be verified

[2]Note that $\delta_3 = 0$.

[3]This construction is inspired by the functional representation lemma.

[4]This is known as Gallager's mapping [26], which is widely used to construct source or channel codes with non-uniform empirical distribution [27], [28].

[5]The representation of such backward channels can be viewed as a manifestation of the functional representation lemma. Moreover, it is instructive to view $\phi_i(\mathcal{C}_i')$ as a codebook generated by V_i'.

that

$$
\begin{aligned}
&I(Y_2;W_2|W_1)\\
&= I(V_1'\oplus Z_1'\oplus N_1'\oplus N_2; V_1'\oplus Z_1'\oplus N_1'\oplus N_2\oplus Z_2\oplus V_2)\\
&= H(V_1'\oplus Z_1'\oplus N_1'\oplus N_2\oplus Z_2\oplus V_2) - H(Z_2\oplus V_2)\\
&= h_b(\delta_1 * d_1 * p_1 * p_2 * d_2 * \delta_2) - h_b(d_2 * \delta_2),\\
&I(W_1;W_2) = 1 - H(V_1'\oplus Z_1'\oplus N_1'\oplus N_2\oplus Z_2\oplus V_2)\\
&= 1 - h_b(\delta_1 * d_1 * p_1 * p_2 * d_2 * \delta_2),\\
&I(W_1,W_2;U_3) = H(W_1,W_2) - H(W_1,W_2|U_3)\\
&= 1 + h_b(\delta_1 * d_1 * p_1 * p_2 * d_2 * \delta_2)\\
&- H(Z_3'\oplus N_3'\oplus N_1\oplus Z_1\oplus V_1, Z_3'\oplus N_3'\oplus N_2\oplus Z_2\oplus V_2)\\
&= H(W_1,W_2) - H(W_1,W_2,U_3) + 1,\\
&I(Y_3;U_3|W_1,W_2) = I(Y_3;U_3) - I(W_1,W_2;U_3)\\
&= 1 - h_b(d_3) - I(W_1,W_2;U_3). \qquad (21)
\end{aligned}
$$

4) For $i \in \mathcal{I}_2$, to partition $\mathcal{C}_i'$ into $2^{n[I(Y_i;U_i|W_1,\cdots,W_i,U_{i+1},\cdots,U_3)+\epsilon_{i,4}]}$ bins[6] with each bin containing $2^{n[I(W_1,\cdots,W_{i-1},U_{i+1},\cdots,U_3;U_i|W_i)-\epsilon_{i,4}]}$ codewords, an LDPC code is used with $2^{n(I(W_1,\cdots,W_{i-1},U_{i+1},\cdots,U_3;U_i|W_i)-\epsilon_{i,4})}$ codewords, each of length nK_i. The parity-check matrix of this LDPC code is $\mathbf{H}_1 = (\tilde{\mathbf{H}}_1, \hat{\mathbf{H}}_1)$ for $i = 1$ and $\mathbf{H}_2' = (\tilde{\mathbf{H}}_2', \hat{\mathbf{H}}_2')$ for $i = 2$, where $\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_2'$ are the parity-check matrices of $\mathcal{C}_1'$ and $\mathcal{C}_2'$, respectively. Thus, we have

$$
\begin{aligned}
&I(U_2,U_3;U_1|W_1)\\
&= I(V_1'\oplus Z_1'\oplus N_1'\oplus N_2\oplus Z_2, V_1'\oplus Z_1'\\
&\qquad \oplus N_1'\oplus N_3\oplus Z_3; V_1')\\
&= H(V_1'\oplus Z_1'\oplus N_1'\oplus N_2\oplus Z_2, V_1'\oplus Z_1'\\
&\qquad \oplus N_1'\oplus N_3\oplus Z_3)\\
&\quad - H(Z_1'\oplus N_1'\oplus N_2\oplus Z_2, Z_1'\oplus N_1'\oplus N_3\oplus Z_3),\\
&I(Y_1;U_1|W_1,U_2,U_3) = I(Y_1;U_1|W_1)\\
&\quad - I(U_2,U_3;U_1|W_1),\\
&I(W_1,U_3;U_2|W_2)\\
&= I(V_2'\oplus Z_2'\oplus N_2'\oplus N_1\oplus Z_1\\
&\qquad \oplus V_1, V_2'\oplus Z_2'\oplus N_2'\oplus N_3\oplus Z_3; V_2')\\
&= H(V_2'\oplus Z_2'\oplus N_2'\oplus N_1\oplus Z_1\\
&\qquad \oplus V_1, V_2'\oplus Z_2'\oplus N_2'\oplus N_3\oplus Z_3)\\
&\quad - H(Z_2'\oplus N_2'\oplus N_1\oplus Z_1\oplus V_1, Z_2'\\
&\qquad \oplus N_2'\oplus N_3\oplus Z_3),\\
&I(Y_2;U_2|W_1,W_2,U_3)\\
&= I(Y_2;U_2|W_2) - I(W_1,U_3;U_2|W_2). \qquad (22)
\end{aligned}
$$

Encoding: Different from the information-theoretic description in Section III-A, we shall interpret joint typicality encoding as the minimum Hamming distance encoding, which is then implemented using the BiP algorithm.

1) For $l \in \mathcal{I}_3$ and a given y_l^n, the l-th encoder finds a codeword $w_l^n \in \mathcal{C}_{W_l}$ from an LDGM code that is the closest (in the Hamming distance) to y_l^n.

[6]This induces a partition of $\mathcal{C}_{U_i}(w_i^n)$.

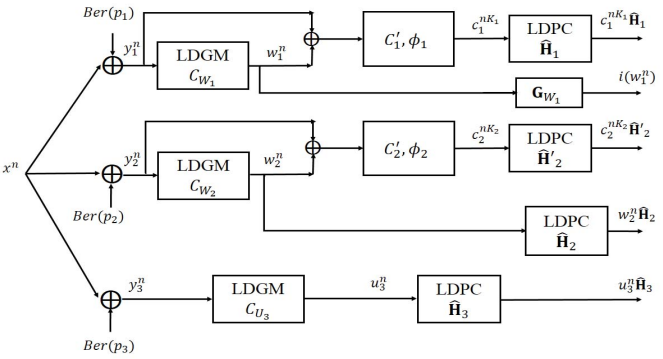


Fig. 2. The proposed encoding scheme.

2) For $i \in \mathcal{I}_2$, find a codeword $c_i^{nK_i} \in \mathcal{C}_i'$ such that $\phi_i(c_i^{nK_i})$ is the closest (in the Hamming distance) to $y_i^n \oplus w_i^n$.
3) Send the index of w_1^n and the syndrome $c_1^{nK_1}\hat{\mathbf{H}}_1$ from the first link to the decoder; note that $w_1^n = i(w_1^n)\mathbf{G}_{W_1}$, where $\mathbf{G}_{W_1}$ is the generator matrix of LDGM code $\mathcal{C}_{W_1}$. Also, send the syndromes $w_2^n\hat{\mathbf{H}}_2$ and $c_2^{nK_2}\hat{\mathbf{H}}_2'$ from the second link to the decoder. Finally, send the syndrome $u_3^n\hat{\mathbf{H}}_3$ from the third link to the decoder.

The block diagram of the proposed encoding scheme is depicted in Fig. 2.

Decoding: Different from the information-theoretic description in Section III-A, we shall interpret joint typicality decoding as maximum *a posteriori* decoding, which is then implemented using the SP algorithm.

1) The decoder first sets $\hat{w}_1^n = w_1^n$.
2) It then finds the most likely choice of w_2^n, denoted by $\hat{w}_2^n$, based on $\hat{w}_1^n$ and $w_2^n\mathbf{H}_2$ (which can be deduced from $w_2^n\hat{\mathbf{H}}_2$ and the fact that $w_2^n\tilde{\mathbf{H}}_2$ is a zero vector). This can be realized via conventional Slepian-Wolf decoding with $\mathbf{H}_2$ defining the factor graph and $\hat{w}_1^n$ serving as side information (see, e.g., [29]).
3) It then finds the most likely choice of u_3^n, denoted by $\hat{u}_3^n$, based on $\hat{w}_1^n$, $\hat{w}_2^n$, and $u_3^n\mathbf{H}_3$ (which can be deduced from $u_3^n\hat{\mathbf{H}}_3$ and the fact that $u_3^n\tilde{\mathbf{H}}_3$ is a zero vector). This can be realized via conventional Slepian-Wolf decoding with $\mathbf{H}_3$ defining the factor graph and $(\hat{w}_1^n, \hat{w}_2^n)$ serving as side information.
4) It then finds the most likely choice of $c_2^{nK_2}$, denoted by $\hat{c}_2^{nK_2}$, based on $\hat{w}_1^n$, $\hat{w}_2^n$, $\hat{u}_3^n$, and $c_2^{nK_2}\mathbf{H}_2'$ (which can be deduced from $c_2^{nK_2}\hat{\mathbf{H}}_2'$ and the fact that $c_2^{nK_2}\tilde{\mathbf{H}}_2'$ is a zero vector). This can be realized via joint demapping and decoding with $(\mathbf{H}_2', \phi_2)$ defining the factor graph and $(\hat{w}_1^n, \hat{w}_2^n, \hat{u}_3^n)$ serving as the channel output (see, e.g., [30]). Set $\hat{u}_2^n = \hat{w}_2^n \oplus \phi_2(\hat{c}_2^{nK_2})$.
5) It then finds the most likely $c_1^{nK_1}$, denoted by $\hat{c}_1^{nK_1}$, based on $\hat{w}_1^n$, $\hat{u}_2^n$, $\hat{u}_3^n$, and $c_1^{nK_1}\mathbf{H}_1$ (which can be deduced from $c_1^{nK_1}\hat{\mathbf{H}}_1$ and the fact that $c_1^{nK_1}\tilde{\mathbf{H}}_1$ is a zero vector). This can be realized via joint demapping and decoding with $(\mathbf{H}_1, \phi_1)$ defining the factor graph

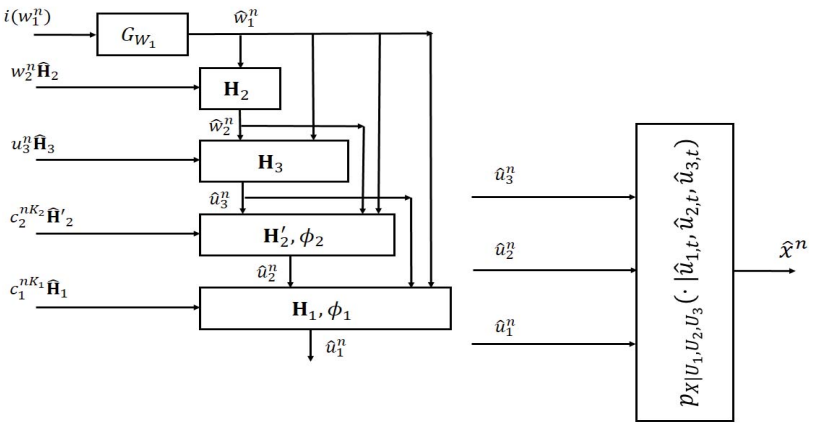


Fig. 3. The proposed successive decoding scheme.

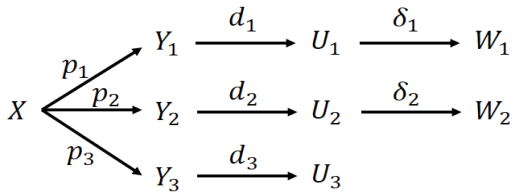


Fig. 4. Joint distribution diagram of the binary CEO problem.

and $(\hat{w}_1^n, \hat{u}_2^n, \hat{u}_3^n)$ serving as the channel output. Set $\hat{u}_1^n = \hat{w}_1^n \oplus \phi_1(\hat{c}_1^{nK_1})$.

6) Finally, it produces a soft reconstruction $\hat{x}^n$ based on $\hat{u}_1^n$, $\hat{u}_2^n$, and $\hat{u}_3^n$ (see (16)).

The block diagram of the proposed decoding scheme is depicted in Fig. 3.

C. Analysis of the Proposed Coding Scheme

Now we proceed to specify the sizes of generator matrices and parity-check matrices used in the proposed scheme and other relevant parameters, assuming that d_1, d_2, d_3, δ_1, and δ_2 are given according to the joint distribution of a 3-link binary CEO problem depicted in Fig. 4.

For the LDGM codes $\mathcal{C}_{W_l}$, as shown in Fig. 2, their generator matrices are of size $m_l \times n$, $l \in \mathcal{I}_3$, respectively, where

$$\begin{aligned} \frac{m_i}{n} &= 1 - h_b(d_i * \delta_i) + \epsilon_{i,1}, \quad i \in \mathcal{I}_2, \\ \frac{m_3}{n} &= 1 - h_b(d_3) + \epsilon_{3,1}. \end{aligned} \tag{23}$$

Furthermore, size of the generator matrix of the LDGM code $\mathcal{C}'_i$ is $M_i \times nK_i$, for $i \in \mathcal{I}_2$. By properly designing these LDGM codes and increasing the block length n, one can ensure that

$$\begin{aligned} &\mathbb{E}\Big(\frac{1}{n}\sum_{j=1}^{n}[y_{i,j} \oplus w_{i,j}]\Big) \approx d_i * \delta_i, \quad i \in \mathcal{I}_2, \\ &\mathbb{E}\Big(\frac{1}{n}\sum_{j=1}^{n}[y_{l,j} \oplus u_{l,j}]\Big) \approx d_l, \quad \epsilon_{l,1} \approx 0,\ l \in \mathcal{I}_3. \end{aligned} \tag{24}$$

For the LDPC codes shown in Fig. 2, the sizes of their parity-check matrices are given as follows:

$$\begin{aligned} &\mathbf{H}_1\colon\ nK_1 \times (nK_1 - M_1 + k_1), \quad \mathbf{H}_2\colon\ n \times (n - m_2 + k_2), \\ &\mathbf{H}'_2\colon\ nK_2 \times (nK_2 - M_2 + k'_2), \quad \mathbf{H}_3\colon\ n \times (n - m_3 + k_3). \end{aligned} \tag{25}$$

All of the LDPC codes in the decoder performs an SP algorithm. Basically, each SP algorithm is an iterative message-passing algorithm which passes LLR values between variable nodes and check nodes of the LDPC code. In each iteration, we have LLR-updating equations in both variable nodes and check nodes. Generally, there are two types of inputs in this algorithm: (1) Syndrome, (2) Side information. In the SP algorithm, initial LLR values for the variable nodes are calculated based on the joint distribution diagram parameters. For instance, the initial LLR values of the SP algorithm by using parity-check matrix $\mathbf{H}_3$ are as follows:

$$\mathrm{LLR}_{t,0} = \log \frac{p_{U_3|W_1,W_2}(0|\hat{w}_{1,t}, \hat{w}_{2,t})}{p_{U_3|W_1,W_2}(1|\hat{w}_{1,t}, \hat{w}_{2,t})}, \quad \text{for } t \in \mathcal{I}_n \ \text{ or } t \in \mathcal{I}_{nK_i}. \tag{26}$$

There are four possible cases for the $\mathrm{LLR}_{t,0}$ based on the values of $(\hat{w}_{1,t}, \hat{w}_{2,t}) \in \mathbb{B}^2$, and all of them can be calculated from the joint distribution diagram.

In the syndrome-decoding part of our proposed scheme, which is implemented by successive SP algorithms, if the optimized degree distributions for the BSC are used with sufficiently long LDPC codes, the Bit Error Rate (BER) for the reconstruction of $\{U_1, U_2, U_3\}$ can be made very close to zero, i.e., $\mathrm{BER}_l \approx 0$ for $l \in \mathcal{I}_3$. In such a case, the total distortion of the l-th link approximately equals d_l. In designing procedure of LDPC codes that are employed for the syndrome-generation and the syndrome-decoding, the following relations are considered in their code rates,

$$\begin{aligned} \mathbf{H}_2\colon\ \frac{m_2 - k_2}{n} &= I(W_1; W_2) - \epsilon_{2,3} \\ &= 1 - h_b(P_1 * \delta_1 * P_2 * \delta_2) - \epsilon_{2,3}, \\ \mathbf{H}_3\colon\ \frac{m_3 - k_3}{n} &= I(W_1, W_2; U_3) - \epsilon_{3,3} \end{aligned}$$

$$\begin{aligned}
&= 2 + h_b(P_1 * \delta_1 * P_2 * \delta_2) \\
&\quad - H(W_1, W_2, U_3) - \epsilon_{3,3}, \\
\mathbf{H}'_2: \ \frac{M_2 - k'_2}{nK_2} &= I(W_1, U_3; U_2|W_2) - \epsilon_{2,4} \\
&= H(W_1, W_2, U_3) - H(W_1, U_2, U_3) - \epsilon_{2,4}, \\
\mathbf{H}_1: \ \frac{M_1 - k_1}{nK_1} &= I(U_2, U_3; U_1|W_1) - \epsilon_{1,4} \\
&= H(W_1, U_2, U_3) - H(U_1, U_2, U_3) - \epsilon_{1,4},
\end{aligned} \tag{27}$$

where $P_l = p_l * d_l$ for $l \in \mathcal{I}_3$. Note that, there are four compound LDGM-LDPC codes[7] in the proposed scheme for a 3-link binary CEO problem. They comprise the LDGM codes $\mathcal{C}_{W_2}$, $\mathcal{C}_{U_3}$, $\mathcal{C}'_1$, and $\mathcal{C}'_2$, respectively with the LDPC codes of parity-check matrices $\mathbf{H}_2$, $\mathbf{H}_3$, $\mathbf{H}_1$, and $\mathbf{H}'_2$.

IV. Analytical Results

It is clear that, for the proposed scheme, there is freedom in choosing $(d_1, \cdots, d_L)$ and $(\delta_1, \cdots, \delta_L)$. The role of $(d_1, \cdots, d_L)$ is to specify the dominant face $\mathcal{F}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$ (and consequently the sum rate) while the role of $(\delta_1, \cdots, \delta_L)$ is to specify the location of the target rate tuple $(R_1, \cdots, R_L)$ on the dominant face. Note that for any $(R_1, \cdots, R_L, D) \in \mathcal{RD}_{\text{CEO}}(p_{U_l|Y_l}, l \in \mathcal{I}_L)$, we have $\sum_{l=1}^{L} R_l \geq R_{\text{th}}$ and $D \geq D_{\text{th}}$, where $R_{\text{th}} = I(Y_{\mathcal{I}_L}; U_{\mathcal{I}_L})$, and $D_{\text{th}} = H(X|U_{\mathcal{I}_L})$. One can interpret R_{th} and D_{th} as the minimum achievable sum rate and distortion associated with a given $(d_1, \cdots, d_L)$. Therefore, it is natural to choose $(d_1, \cdots, d_L)$ that achieves an optimal tradeoff between R_{th} and D_{th}, which motivates the following definition.

Definition 5: An L-tuple $(d_1^*, \cdots, d_L^*)$ is called an optimal d-allocation if it is a minimizer of F for a certain $\mu \geq 0$, where

$$F = D_{\text{th}} + \mu R_{\text{th}}. \tag{28}$$

We shall derive several analytical results surrounding Definition 5. An investigation along this line was initiated in [20] for the case $L = 2$.

Note that

$$\begin{aligned}
R_{\text{th}} &= H(U_{\mathcal{I}_L}) - H(U_{\mathcal{I}_L}|Y_{\mathcal{I}_L}) \\
&= H(U_{\mathcal{I}_L}) - H(Z_{\mathcal{I}_L}) \\
&= H(U_{\mathcal{I}_L}) - \sum_{l=1}^{L} h_b(d_l),
\end{aligned}$$

and

$$\begin{aligned}
D_{\text{th}} &= H(X, U_{\mathcal{I}_L}) - H(U_{\mathcal{I}_L}) \\
&= H(X) + H(U_{\mathcal{I}_L}|X) - H(U_{\mathcal{I}_L}) \\
&= H(X) + H(N_{\mathcal{I}_L} \oplus Z_{\mathcal{I}_L}) - H(U_{\mathcal{I}_L}) \\
&= 1 + \sum_{l=1}^{L} h_b(P_l) - H(U_{\mathcal{I}_L}),
\end{aligned}$$

where $P_l = p_l * d_l$, $l \in \mathcal{I}_L$. Define $Q_j = \prod_{l=1}^{L} \eta(P_l, b_l(j))$ for $j \in [0 : 2^L - 1]$, where $b_l(j)$ denotes the l-th digit in the binary expansion of j, and

$$\eta(P_l, b_l(j)) = \begin{cases} P_l, & b_l(j) = 0, \\ 1 - P_l, & b_l(j) = 1. \end{cases}$$

For example, when $L = 3$, we have

$$\begin{aligned}
Q_0 &= P_1P_2P_3, & Q_4 &= (1-P_1)P_2P_3, \\
Q_1 &= P_1P_2(1-P_3), & Q_5 &= (1-P_1)P_2(1-P_3), \\
Q_2 &= P_1(1-P_2)P_3, & Q_6 &= (1-P_1)(1-P_2)P_3, \\
Q_3 &= P_1(1-P_2)(1-P_3), & Q_7 &= (1-P_1)(1-P_2)(1-P_3).
\end{aligned} \tag{29}$$

It can be verified that

$$H(U_{\mathcal{I}_L}) = -\sum_{j=0}^{2^L-1} \left[\frac{Q_j + Q_{2^L-1-j}}{2}\right] \log \left[\frac{Q_j + Q_{2^L-1-j}}{2}\right]. \tag{30}$$

Lemma 1: For the objective function F defined in (28), its minimum value is equal to 1 when $\mu \geq 1$.

Proof: It is clear that $F = 1$ when $(d_1, \cdots, d_L) = (0.5, \cdots, 0.5)$. Now assume that a certain choice of $(d_1, \cdots, d_L)$ gives $F < 1$. As a consequence, we have

$$\mu < \frac{1 - D_{\text{th}}}{R_{\text{th}}} = \frac{H(U_{\mathcal{I}_L}) - \sum_{l=1}^{L} h_b(p_l * d_l)}{H(U_{\mathcal{I}_L}) - \sum_{l=1}^{L} h_b(d_l)} \leq 1, \tag{31}$$

which is contradictory with the fact that $\mu \geq 1$. □

Lemma 2: Let $p_1 \leq p_2$ and $d_1 > d_2$. If $P_1 = p_1 * d_1$, $P_2 = p_2 * d_2$, $P'_1 = p_1 * d_2$, and $P'_2 = p_2 * d_1$, then

$$\begin{aligned}
P_1 + P_2 &> P'_1 + P'_2, \\
P_1P_2 &> P'_1P'_2, \\
2[P_1P_2 - P'_1P'_2] &= [P_1 + P_2] - [P'_1 + P'_2].
\end{aligned} \tag{32}$$

Proof: The proof is straightforward. □

Lemma 3: Let $p_1 \leq p_2 \leq \cdots \leq p_L$. If $d_1 > d_2$ in the L-tuple $(d_1, \cdots, d_L)$, then by swapping d_1 and d_2, the value of $H(U_{\mathcal{I}_L})$ will increase.

Proof: See Appendix A.

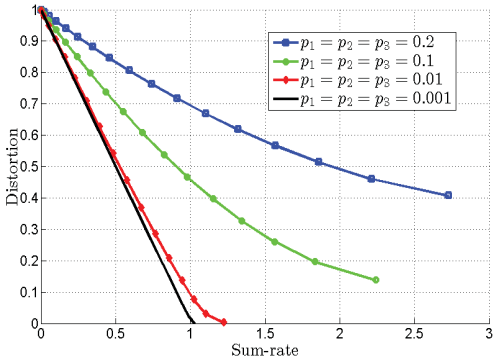
Lemma 4: If $p_1 \leq p_2 \leq \cdots \leq p_L$, then $d_1^* \leq d_2^* \leq \cdots \leq d_L^*$.

Proof: Assume that this is not true, and thus there exits i such that $d_i^* > d_{i+1}^*$. We prove that by swapping d_i^* and d_{i+1}^*, the objective function $F = D_{\text{th}} + \mu R_{\text{th}}$ will decrease, which is a contradiction. Note that

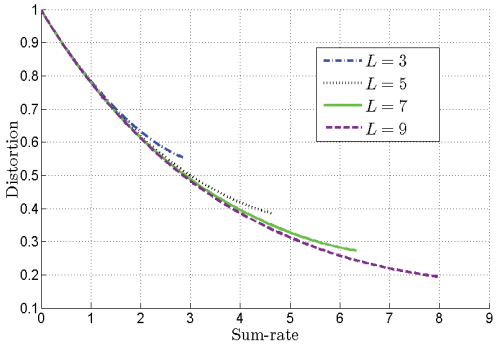
$$F = 1 + \sum_{l=1}^{L} h_b(p_l * d_l^*) - \mu \sum_{l=1}^{L} h_b(d_l^*) + (\mu - 1) H(U_{\mathcal{I}_L}). \tag{33}$$

Based on Lemmas 1 and 2, term $(\mu - 1)H(U_{\mathcal{I}_L})$ decreases by swapping d_i^* and d_{i+1}^*. Also, the term $-\mu \sum_{l=1}^{L} h_b(d_l^*)$ clearly remains unchanged by this replacement. Without loss

[7] Generally, there is a compound LDGM-LDPC code in the first and the L-th link; and there are two compound codes in the i-th link, for $i \in [2 : L-1]$. Thus, there are totally $2L - 2$ compound codes in an L-link case.



(a) Equal noise parameters with the same quantizer in each link, the number of links is $L = 3$.



(b) $p_l = 0.25$ for $l \in \mathcal{I}_L$.

Fig. 5. Sum-rate vs. distortion curves under the BSC assumption for the test channels.

of generality, let us assume $i = 1$. Therefore, it is enough to show that

$$h_b(p_1 * d_2^*) + h_b(p_2 * d_1^*) < h_b(p_1 * d_1^*) + h_b(p_2 * d_2^*). \tag{34}$$

By defining the following variables z_1 and z_2, we have:

$$\begin{aligned} z_1 &\triangleq p_1 * d_1^* \Rightarrow p_1 * d_2^* < z_1 < p_2 * d_1^*, \\ z_2 &\triangleq p_1 * d_2^* + p_2 * d_1^* - p_1 * d_1^* \\ &\Rightarrow p_1 * d_2^* < z_2 < p_2 * d_2^* < p_2 * d_1^*, \\ &\Rightarrow z_1 + z_2 = p_1 * d_2^* + p_2 * d_1^*. \end{aligned} \tag{35}$$

Since, $h_b(x)$ is a concave function in x, from (35)

$$h_b(p_1 * d_2^*) + h_b(p_2 * d_1^*) < h_b(z_1) + h_b(z_2). \tag{36}$$

Furthermore, $h_b(x)$ is an increasing function in the interval $[0, 0.5]$. Thus,

$$h_b(z_1) + h_b(z_2) < h_b(p_1 * d_1^*) + h_b(p_2 * d_2^*). \tag{37}$$

From (36) and (37), the inequality (34) is concluded. Hence, the proof is completed. □

V. Numerical Results

Now we provide some numerical examples of optimal d-allocations. Without loss of generality, we assume $p_1 \leq p_2 \leq \cdots \leq p_L$. It follows by Lemma 4 that $d_1^* \leq d_2^* \leq \cdots \leq d_L^*$ for the resulting optimal d-allocation. Obviously, d_l^* equals 0 for all l's when $\mu = 0$. There exists a $\mu_0 > 0$ such that for $0 \leq \mu < \mu_0$, all L links are involved in information sending, i.e., $d_l^* < 0.5$ for $l \in \mathcal{I}_L$, while $d_L^* = 0.5$ for $\mu = \mu_0$. Therefore, the L-th link becomes inactive for $\mu \geq \mu_0$. Accordingly, the problem is reduced to an $(L-1)$-link case. By increasing μ, the noisy links are eliminated one-by-one. Finally, it is reduced to the case of $L = 2$. We illustrate this phenomenon through the following simple example.

Example 1: Let $L = 3$ and $p_1 = p_2 = p_3 = 0.1$. Based on the numerical results, if $0 \leq \mu < \mu_0 \approx 0.3923$, then the straight line $0 \leq d_1^* = d_2^* = d_3^* < 0.125$ determines the location of the optimal points. For $\mu_0 \leq \mu < \mu_1 \approx 0.42$, we have $d_1^* = d_2^* \leq 0.125$ and $0.125 < d_3^* < 0.5$. If $\mu = \mu_1$, then $d_1^* = d_2^* = 0.089$ and $d_3^* = 0.5$. Similarly, if $\mu_1 < \mu < \mu_2 \approx 0.4245$, then $d_1^* < d_2^* < 0.5$ and $d_3^* = 0.5$. Next, for $\mu_2 \leq \mu < \mu_{\max} = 0.64$, the first link is only involved in sending the information, i.e., $0.023 < d_1^* < 0.5$ and $d_2^* = d_3^* = 0.5$. Finally, $d_1^* = d_2^* = d_3^* = 0.5$ for $\mu \geq \mu_{\max}$.

The next example illustrates the sum-rate-distortion tradeoffs under equal d-allocation (i.e., $d_1 = d_2 = \cdots = d_L$).

Example 2: Let $L = 3$ and $p_1 = p_2 = p_3$. The sum-rate distortion curves under equal d-allocation are depicted in Fig. 5(a) for various noise parameters. In Fig. 5(b), the sum-rate distortion curves under equal d-allocation are shown for the case of $p_l = 0.25$ with $L = 3, 5, 7, 9$.

Example 3: Based on the numerical and the analytical results presented in [20], for a two-link binary CEO problem, the equal allocation, i.e., $d_1^* = d_2^*$, is not an optimal d-allocation for some values of sum-rate and distortion, even in the case of equal noise parameters $p_1 = p_2$. Here, it is shown that this surprising result is also authentic for the multi-link case. In Fig. 6, the sum-rate distortion curves are shown for some cases. As it is seen, involving all the links does not necessarily provide minimum values of the sum-rate and the distortion.

Example 4: In this example, a 3-link binary CEO problem is considered with almost prominent differences between the values of the noise parameters. As an example, let $p_1 = 0.01$, $p_2 = 0.1$, and $p_3 = 0.2$. The sum-rate versus the distortion curves are presented in Fig. 7. It is assumed that the binary quantizers in each link are the same, when more than one link are involved in sending the information. Clearly, utilizing low noise links provides better results.

Now we proceed to present some experimental results for the proposed coding scheme. In our implementation, the degree distributions of the LDPC codes are provided in Appendix B; furthermore, the degree distributions of the LDGM codes are designed based on the method proposed in [16], where the degrees of check nodes are regular and those of variable nodes follow a Poisson distribution. The relevant parameters of the proposed scheme are presented in Tables II and III. In particular, each choice of $(d_1, \cdots, d_L)$

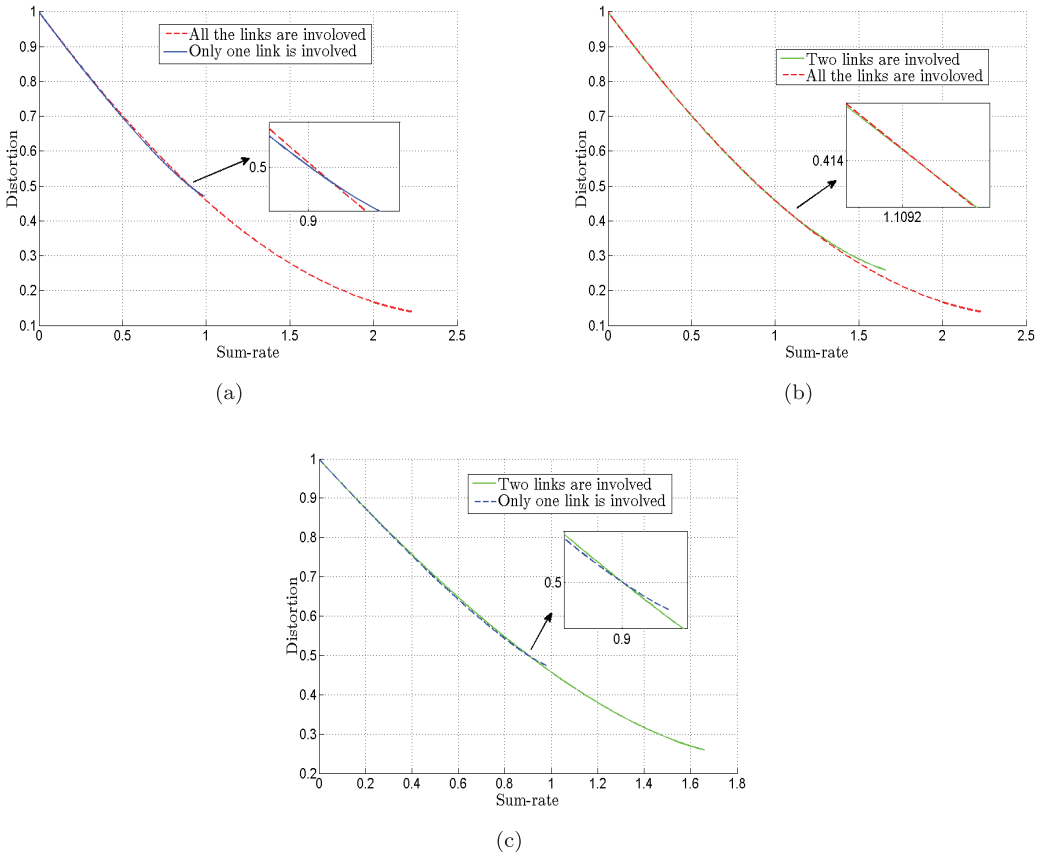


Fig. 6. Sum-rate vs. distortion curves under the BSC assumption for the test channels. $L = 3$ and $p_1 = p_2 = p_3 = 0.1$ in different allocation scenarios.

TABLE II

PARAMETERS AND NUMERICAL RESULTS OF THE PROPOSED CODING SCHEME, (EXAMPLE 5)

m_1, m_2, m_3	k_1, k_2, k_2', k_3	δ_1, δ_2	d_1, d_2, d_3	R_1, R_2, R_3, R	BER_1,BER_2,BER_3,BER_4	D_{em}	Gap
5400, 4000, 500	–, 3700, –, 500	0, 0	0.102, 0.1669, 0.3783	0.54, 0.4, 0.05, 0.99	0.0022, 0.0025, –, –	0.7532	0.0289
3300, 2100, 500	9900, 2000, 9950, 500	0.1024, 0.141	0.1036, 0.1677, 0.3783	0.538, 0.3868, 0.05, 0.9748	0.001, 0.0014, 0.0014, 0.0021	0.759	0.0347
26500, 18250, 2300	–, 18000, –, 2300	0, 0	0.1014, 0.1658, 0.3778	0.53, 0.36, 0.046, 0.936	0.0012, 0.0.0016, –, –	0.7474	0.0231
16000, 9500, 2300	49500, 9500, 49750, 2300	0.1024, 0.141	0.1019, 0.166, 0.3778	0.5304, 0.3677, 0.0460, 0.9441	0.0009, 0.0.0012, 0.001, 0.0015	0.7552	0.0309
52500, 36000, 4500	–, 35000, –, 4500	0, 0	0.1009, 0.1653, 0.3776	0.525, 0.35, 0.045, 0.92	0.001, 0.0013, –, –	0.7438	0.0195
31500, 18000, 4500	99000, 17500, 99500, 4500	0.1024, 0.141	0.1014, 0.1648, 0.3776	0.5261, 0.3542, 0.045, 0.9253	0.0006, 0.0011, 0.0008, 0.0014	0.7503	0.026

TABLE III

PARAMETERS AND NUMERICAL RESULTS OF THE PROPOSED CODING SCHEME, (EXAMPLE 6)

$m_{1\le i\le 3}, m_4$	$k_1, k_2, k_2', k_3, k_3', k_4$	$\delta_1, \delta_2, \delta_3$	d_1, d_2, d_3, d_4	R_1, R_2, R_3, R_4, R	BER_1,BER_2,BER_3,BER_4,BER_5,BER_6	D_{em}	Gap
5500, 5500	–, 4400, –, 4000, –, 3600	0, 0, 0	0.1025, 0.1028, 0.1031, 0.1022	0.55, 0.44, 0.4, 0.36, 1.75	0.0024, 0.0027, 0.003, –, –, –	0.2743	0.0209
5200, 5500	9950, 4300, 9950, 3800, 9950, 3700	0.01, 0.01, 0.01	0.103, 0.1029, 0.1033, 0.1022	0.5441, 0.4542, 0.4041, 0.37, 1.7724	0.0015, 0.0019, 0.0019, 0.0026, 0.0026, 0.0022	0.2754	0.022
27000, 27000	–, 21500, –, 19000, –, 17000	0, 0, 0	0.102, 0.1021, 0.1025, 0.1019	0.54, 0.43, 0.38, 0.34, 1.69	0.002, 0.002, 0.0023, –, –, –	0.2678	0.0144
25500, 27000	49600, 20500, 49600, 18000, 49600, 17000	0.01, 0.01, 0.01	0.1019, 0.1023, 0.1025, 0.1019	0.5343, 0.4342, 0.3842, 0.34, 1.6927	0.0016, 0.0014, 0.002, 0.0022, 0.0015, 0.0018	0.2694	0.016
53000, 53000	–, 42000, –, 37000, –, 32000	0, 0, 0	0.1017, 0.1019, 0.102, 0.1017	0.53, 0.42, 0.37, 0.32, 1.64	0.0011, 0.0015, 0.0012, –, –, –	0.2629	0.0095
50500, 53000	99100, 40000, 99100, 34000, 99100, 32000	0.01, 0.01, 0.01	0.1015, 0.101, 0.1014, 0.1017	0.5293, 0.4244, 0.3643, 0.32, 1.638	0.0009, 0.0012, 0.001, 0.0013, 0.001, 0.0015	0.2637	0.0103

corresponds to an optimal d-allocation. The rate of each encoder is calculated as follows:

$$\begin{aligned} R_1 &= \big(\frac{m_1}{n}\big) + \big(I(Y_1;U_1|W_1)\big) \times \frac{k_1}{n} \\ &\approx \big(1 - h_b(d_1 * \delta_1)\big) + \big(h_b(d_1 * \delta_1) - h_b(d_1)\big) \times \frac{k_1}{n}, \\ R_l &= \big(\frac{k_l}{n}\big) + \big(I(Y_l;U_l|W_l)\big) \times \frac{k_l'}{n} \\ &= \big(\frac{m_l}{n} - \frac{m_l - k_l}{n}\big) + \big(I(Y_l;U_l|W_l)\big) \times \frac{k_l'}{n} \\ &\approx \big(1 - h_b(d_l * \delta_l)\big) - \big(I(W_1, \cdots, W_{l-1}; W_l)\big) \\ &\quad + \big(h_b(d_l * \delta_l) - h_b(d_l)\big) \times \frac{k_l'}{n}, \quad l \in [2 : L-1], \\ R_L &= \big(\frac{k_L}{n}\big) = \big(\frac{m_L}{n}\big) - \big(\frac{m_L - k_L}{n}\big) \\ &\approx \big(1 - h_b(d_L)\big) - \big(I(W_1, \cdots, W_{L-1}; U_L)\big). \end{aligned} \tag{38}$$

Example 5: Consider a 3-link case. Let $p_1 = 0.2$, $p_2 = 0.205$, and $p_3 = 0.21$ as well. For $\mu = 0.25$, the optimal d-allocation is given by $d_1^* = 0.1$, $d_2^* = 0.164$, and $d_3^* = 0.377$; as consequence, we have $R_{\mathrm{th}} = 0.9091$ and $D_{\mathrm{th}} = 0.7243$. The performance of the proposed coding scheme is presented for the corner and the intermediate points separately.

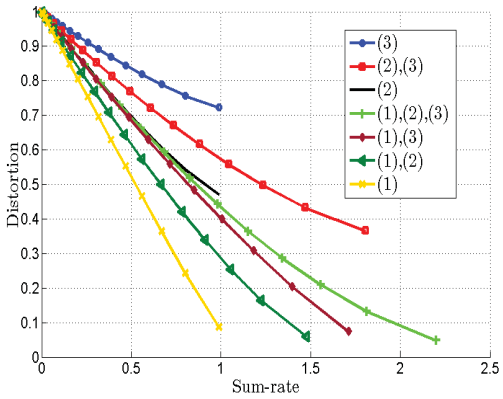


Fig. 7. Sum-rate vs. distortion curves under the BSC assumption for the test channels (Example 4). The number of involved links are given in the legend.

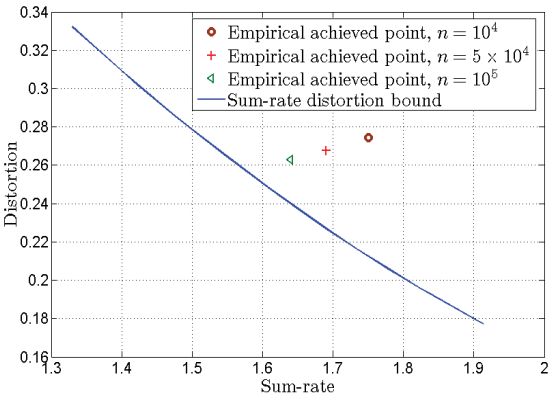


Fig. 9. Performance of the sum-rate vs. distortion for the implemented codes, (Example 6).

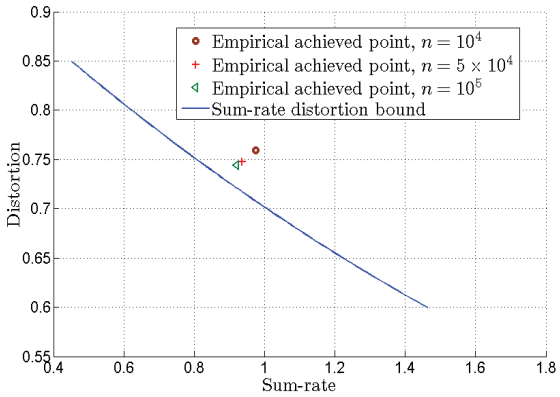


Fig. 8. Performance of the sum-rate vs. the distortion for the implemented codes, (Example 5).

The block lengths are equal to $n = 10^4$, $n = 5 \times 10^4$, and $n = 10^5$. First, to achieve a corner point, we set $\delta_1 = \delta_2 = 0$. However, in order to achieve an intermediate point, any choice of $\delta_1 \in (0, 0.5)$ and $\delta_2 \in (0, 0.5)$ gives a specific intermediate point on the dominant face. In this example, we set $K_1 = 7$, $K_2 = 6$, $M_1 = 0.22\ nK_1$, and $M_2 = 0.19\ nK_2$ for the intermediate point. The results are presented in Table II and Fig. 8. The gap values for the code lengths $n = 10^4$, 5×10^4, and 10^5 are about 0.029, 0.023, and 0.02, respectively.

Example 6: Consider a 4-link case and let $p_l = 0.1$ for $l \in \mathcal{I}_4$. For $\mu = 0.27$, the optimal d-allocation is given by $d_l^* = 0.1$, for $l \in \mathcal{I}_4$; as a consequence, we have $R_{\text{th}} = 1.591$ and $D_{\text{th}} = 0.2534$. The block lengths are set to $n = 10^4$, $n = 5 \times 10^4$, and $n = 10^5$. In order to achieve a corner point, we set $\delta_1 = \delta_2 = \delta_3 = 0$. However, to achieve an intermediate point, any choice of $\delta_i \in (0, 0.5)$ for $i \in \mathcal{I}_3$, gives a specific intermediate point. In this example, we set $K_i = 9$ and $M_i = 0.12\ nK_i$, $i \in \mathcal{I}_3$, for the intermediate point. The results are shown in Table III and Fig. 9. The gap values for the code lengths $n = 10^4$, 5×10^4, and 10^5 are about 0.021, 0.015, and 0.01, respectively. According to the results of Examples 5 and 6, by decreasing the noise parameter or increasing the number of links, the gap values from the theoretical bounds are reduced. Moreover, larger block length n causes smaller gap values.

VI. Conclusion

We have proposed a practical coding scheme for the binary CEO problem under the log-loss criterion based on the idea of quantization splitting. The underlying methodology is in fact quite general and is applicable to the non-binary case as well. It should be emphasized that, to implement the proposed scheme, one needs to first specify the test channel model for each encoder. In general, it is preferable for the system to operate in a mode that corresponds to a certain boundary point of the rate-distortion region. Identifying the boundary-attaining test channel models is an interesting research problem worthy of further investigation.

Appendix A
Proof of Lemma 3

Since $H(U_{\mathcal{I}_L})$ is a function of P_l for $l \in \mathcal{I}_L$, we shall denote it by $H_P(P_1, \cdots, P_L)$. It suffices to show that

$$H_P(P_1', P_2', P_3, \cdots, P_L) > H_P(P_1, P_2, P_3, \cdots, P_L), \quad (39)$$

where $P_1' = p_1 * d_2$ and $P_2' = p_2 * d_1$. From (30),

$$H_P(P_1, P_2, \cdots, P_L) = -\sum_{j=0}^{2^L-1} q_j \log(q_j), \quad (40)$$

where $q_j = \frac{Q_j + Q_{2^L-1-j}}{2}$. Hence, (39) can be written as follows:

$$-\sum_{j=0}^{2^L-1} q_j' \log(q_j') > -\sum_{j=0}^{2^L-1} q_j \log(q_j). \quad (41)$$

Partition (q_j)'s and (q_j')'s in some groups with four members as follows:

$$q_a = \frac{P_1P_2\Psi + (1-P_1)(1-P_2)\Psi'}{2},$$
$$q_b = \frac{P_1(1-P_2)\Psi + (1-P_1)P_2\Psi'}{2},$$
$$q_c = \frac{(1-P_1)P_2\Psi + P_1(1-P_2)\Psi'}{2},$$

$$q_d = \frac{(1-P_1)(1-P_2)\Psi + P_1P_2\Psi'}{2}, \tag{42}$$

and

$$\begin{aligned} q'_a &= \frac{P'_1P'_2\Psi + (1-P'_1)(1-P'_2)\Psi'}{2}, \\ q'_b &= \frac{P'_1(1-P'_2)\Psi + (1-P'_1)P'_2\Psi'}{2}, \\ q'_c &= \frac{(1-P'_1)P'_2\Psi + P'_1(1-P'_2)\Psi'}{2}, \\ q'_d &= \frac{(1-P'_1)(1-P'_2)\Psi + P'_1P'_2\Psi'}{2}, \end{aligned} \tag{43}$$

where Ψ is an arbitrary product of P_i or $(1-P_i)$, for $i \in [3:L]$, and

$$\Psi' = \frac{P_3 \cdots P_L \times (1-P_3)\cdots(1-P_L)}{\Psi}. \tag{44}$$

Without loss of generality, it can be assumed that $\Psi \leq \Psi'$. Therefore, $q_a > q_d$ and $q'_a > q'_d$. By applying Lemma 2,

$$\begin{aligned} &2(q_a - q'_a) \\ &= \Psi[P_1P_2 - P'_1P'_2] + \Psi'[P_1P_2 - P_1 - P_2 - P'_1P'_2 + P'_1 + P'_2] \\ &= \Psi[P_1P_2 - P'_1P'_2] + \Psi'[0] > 0 \Rightarrow q_a > q'_a. \end{aligned} \tag{45}$$

Similarly, we can show $q'_d > q_d$. Thus, $q_a > q'_a > q'_d > q_d$. Now consider two following cases:

1) $P_1 \geq P_2$:

$$\begin{aligned} P_1 \geq P_2 &\Rightarrow q_c \geq q_b, \quad q'_c \geq q'_b. \\ 2(q_c - q'_c) &= \Psi[P'_1P'_2 - P'_2 + P_2 - P_1P_2] \\ &\quad + \Psi'[P'_1P'_2 - P'_1 + P_1 - P_1P_2] \\ &> \Psi[P'_1P'_2 - P'_2 + P_2 - P_1P_2] \\ &\quad + \Psi[P'_1P'_2 - P'_1 + P_1 - P_1P_2] = 0 \\ &\Rightarrow q_c > q'_c. \end{aligned} \tag{46}$$

Note that in this case, $P'_1\ P'_2 - P'_1 + P_1 - P_1\ P_2 = \frac{P'_2 - P'_1 + P_1 - P_2}{2} \geq 0$. Similarly, we can show $q'_b > q_b$. Thus, $q_c > q'_c > q'_b > q_b$.

2) $P_1 < P_2$:

$$\begin{aligned} P_1 < P_2 &\Rightarrow q_c < q_b,\ q'_c < q'_b. \\ 2(q_c - q'_c) &= \Psi[P'_1P'_2 - P'_2 + P_2 - P_1P_2] \\ &\quad + \Psi'[P'_1P'_2 - P'_1 + P_1 - P_1P_2] \\ &< \Psi'[P'_1P'_2 - P'_2 + P_2 - P_1P_2] \\ &\quad + \Psi'[P'_1P'_2 - P'_1 + P_1 - P_1P_2] = 0 \\ &\Rightarrow q_c < q'_c. \end{aligned} \tag{47}$$

Note that in this case, $P'_1\ P'_2 - P'_2 + P_2 - P_1\ P_2 = \frac{P'_1 - P'_2 + P_2 - P_1}{2} \geq 0$. Similarly, we can show $q_b > q'_b$. Thus, $q_b > q'_b > q'_c > q_c$.

Finally, note that

$$q_a + q_b + q_c + q_d = q'_a + q'_b + q'_c + q'_d = \frac{\Psi + \Psi'}{2}. \tag{48}$$

Due to the concavity of the function $f(x) = -x\log(x)$, it is concluded that

$$\begin{aligned} &-q_a\log(q_a) - q_b\log(q_b) - q_c\log(q_c) - q_d\log(q_d) \\ &< -q'_a\log(q'_a) - q'_b\log(q'_b) - q'_c\log(q'_c) - q'_d\log(q'_d). \end{aligned} \tag{49}$$

By doing a summation over all possible values of Ψ in the mentioned 4-tuple groups, (39) is proved. □

APPENDIX B
DEGREE DISTRIBUTIONS

In example 5, the employed degree distribution of parity-check matrices are as follows, which were obtained based on the degree distributions available in [31].

$\mathbf{H}_2$: $\lambda(x) = 0.4145x + 0.1667x^2 + 0.0571x^4 + 0.0737x^5 + 0.0022x^8 + 0.0118x^9 + 0.0751x^{11} + 0.0575x^{19} + 0.0063x^{26} + 0.0046x^{35} + 0.0171x^{43} + 0.0443x^{62} + 0.051x^{82} + 0.0165x^{99}$, and $\rho(x) = 0.5x^2 + 0.5x^3$.

$\mathbf{H}_3$: $\lambda(x) = 0.2911\ x + 0.19x^2 + 0.0408x^4 + 0.0874x^5 + 0.0074x^6 + 0.1125x^7 + 0.0925x^{15} + 0.0186x^{20} + 0.124x^{32} + 0.016x^{39} + 0.02x^{44}$, and $\rho(x) = x^3$.

$\mathbf{H}_1$: $\lambda(x) = 0.41\ x + 0.1724\ x^2 + 0.0995x^4 + 0.0546x^5 + 0.0379x^6 + 0.0312x^{10} + 0.0288x^{14} + 0.0432x^{16} + 0.0217x^{20} + 0.0385x^{28} + 0.0375x^{50} + 0.0023x^{52} + 0.0158x^{62} + 0.0066x^{71}$, and $\rho(x) = 0.4\ x^2 + 0.6\ x^3$.

$\mathbf{H}'_2$: $\lambda(x) = 0.3424x + 0.165x^2 + 0.12x^4 + 0.0191x^5 + 0.012x^6 + 0.1416x^{10} + 0.0211x^{25} + 0.0202x^{26} + 0.0185x^{34} + 0.0429x^{36} + 0.0133x^{38} + 0.0022x^{39} + 0.0104x^{40} + 0.0704x^{99}$, and $\rho(x) = 0.5x^2 + 0.5x^4$.

In example 6, the employed degree distribution of parity-check matrices are as follows, which were obtained based on the degree distributions available in [31].

$\mathbf{H}_2$: $\lambda(x) = 0.3585x + 0.1664x^2 + 0.0487x^4 + 0.1205x^5 + 0.0006x^6 + 0.04x^{10} + 0.0744x^{13} + 0.0339x^{25} + 0.0076x^{30} + 0.0564x^{34} + 0.0918x^{99}$, and $\rho(x) = x^3$.

$\mathbf{H}_3$: $\lambda(x) = 0.3151x + 0.1902x^2 + 0.0449x^4 + 0.1706x^6 + 0.1405x^{17} + 0.0082x^{37} + 0.044x^{41} + 0.0863x^{66}$, and $\rho(x) = 0.5x^3 + 0.5x^4$.

$\mathbf{H}_4$: $\lambda(x) = 0.292x + 0.174x^2 + 0.0523x^4 + 0.0257x^5 + 0.122x^6 + 0.0218x^8 + 0.021x^{10} + 0.0322x^{14} + 0.1128x^{23} + 0.0328x^{31} + 0.0274x^{44} + 0.0048x^{53} + 0.0126x^{59} + 0.0681x^{99}$, and $\rho(x) = x^4$.

$\mathbf{H}_1$, $\mathbf{H}'_2$, and $\mathbf{H}'_3$: $\lambda(x) = 0.3037x + 0.1731x^2 + 0.0671x^4 + 0.0123x^5 + 0.1341x^6 + 0.0314x^{12} + 0.011x^{14} + 0.0257x^{16} + 0.091x^{19} + 0.04x^{39} + 0.0117x^{51} + 0.0189x^{57} + 0.0112x^{62} + 0.0684x^{76}$, and $\rho(x) = 0.4x^2 + 0.6x^4$.

REFERENCES

[1] M. Nangir, R. Asvadi, M. Ahmadian-Attari, and J. Chen, "Successive Wyner-Ziv coding for the binary CEO problem under log-loss," in *Proc. 29th Biennial Symp. Commun. (BSC)*, Jun. 2018, pp. 1–5.

[2] J. He *et al.*, "A tutorial on lossy forwarding cooperative relaying," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 66–87, 1st Quart., 2019.

[3] Z. Kong, S. A. Aly, and E. Soljanin, "Decentralized coding algorithms for distributed storage in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 2, pp. 261–267, Feb. 2010.

[4] P. L. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms and Applications*. Amsterdam, The Netherlands: Elsevier, 2009.

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

[6] A. Vempaty and L. R. Varshney, "The non-regular CEO problem," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2764–2775, May 2015.

[7] K. Eswaran and M. Gastpar, "Remote source coding under Gaussian noise: Dueling roles of power and entropy power," 2018, *arXiv:1805.06515*. [Online]. Available: https://arxiv.org/abs/1805.06515

[8] D. Seo and L. R. Varshney, "The CEO problem with rth power of difference and logarithmic distortions," 2018, *arXiv:1812.00903*. [Online]. Available: https://arxiv.org/abs/1812.00903

[9] T. A. Courtade and T. Weissman, "Multiterminal source coding under logarithmic loss," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 740–761, Jan. 2014.

[10] Y. Zhou, Y. Xu, W. Yu, and J. Chen, "On the optimal fronthaul compression and decoding strategies for uplink cloud radio access networks," *IEEE Trans. Inf. Theory.*, vol. 62, no. 12, pp. 7402–7418, Dec. 2016.

[11] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: MIT Press, 1963.

[12] M. J. Wainwright, E. Maneva, and E. Martinian, "Lossy source compression using low-density generator matrix codes: Analysis and algorithms," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1351–1368, Mar. 2010.

[13] T. Filler, "Minimizing embedding impact in steganography using low density codes," Ph.D. dissertation, Dept. Elect. Comput. Eng., SUNY Binghamton, Binghamton, NY, USA, 2007.

[14] T. Filler and J. Fridrich, "Binary quantization using belief propagation with decimation over factor graphs of LDGM codes," in *Proc. Allerton Conf. Commun., Control, Comput.*, Sep. 2007, pp. 1–5.

[15] M. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1061–1079, Mar. 2009.

[16] M. Nangir, M. Ahmadian-Attari, and R. Asvadi, "Binary Wyner–Ziv code design based on compound LDGM–LDPC structures," *IET Commun.*, vol. 12, no. 4, pp. 375–383, Mar. 2018.

[17] J. Chen and T. Berger, "Successive Wyner–Ziv coding scheme and its application to the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1586–1603, Apr. 2008.

[18] J. Chen, C. Tian, T. Berger, and S. S. Hemami, "Multiple description quantization via Gram-Schmidt orthogonalization," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5197–5217, Dec. 2006.

[19] X. Zhang, J. Chen, S. B. Wicker, and T. Berger, "Successive coding in multiuser information theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2246–2254, Jun. 2007.

[20] M. Nangir, R. Asvadi, M. Ahmadian-Attari, and J. Chen, "Analysis and code design for the binary CEO problem under logarithmic loss," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6003–6014, Dec. 2018.

[21] B. Rimoldi and R. Urbanke, "Asynchronous Slepian-Wolf coding via source-splitting," in *Proc. IEEE ISIT*, Jun. 1997, p. 271.

[22] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.

[23] A. J. Grant, B. Rimoldi, R. L. Urbanke, and P. A. Whiting, "Rate-splitting multiple access for discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 873–890, Mar. 2001.

[24] J. Wang, J. Chen, L. Zhao, P. Cuff, and H. Permuter, "On the role of the refinement layer in multiple description coding and scalable coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1443–1456, Mar. 2011.

[25] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[26] R. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[27] Z. Sun, M. Shao, J. Chen, K. M. Wong, and X. Wu, "Achieving the rate-distortion bound with low-density generator matrix codes," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1643–1653, Jun. 2010.

[28] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "How to achieve the capacity of asymmetric channels," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3371–3393, May 2018.

[29] J. Chen, D.-K. He, and A. Jagmohan, "The equivalence between slepian-wolf coding and channel coding under density evolution," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2534–2540, Sep. 2009.

[30] J. Cao, S. Hranilovic, and J. Chen, "Capacity and nonuniform signaling for discrete-time Poisson channels," *J. Opt. Commun. Netw.*, vol. 5, no. 4, pp. 329–337, Apr. 2013.

[31] D. H. Schonberg, "Practical distributed source coding and its application to the compression of encrypted data," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, Berkeley, CA, USA, 2007.

Mahdi Nangir received the B.Sc. degree (Hons.) in electrical engineering from the University of Tabriz, Tabriz, Iran, and the M.Sc. degree in communication system engineering from the Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively, and the Ph.D. degree from the K. N. Toosi University of Technology, Tehran, in 2018. In 2017, he joined McMaster University, Hamilton, ON, Canada, as a Research Visiting Student. He is currently an Assistant Professor with the Faculty of Electrical and Computer Engineering, University of Tabriz. His research interests include coding and information theory, distributed source coding, and data compression algorithms. He was a finalist of the Mathematics Olympiad and received a bronze medal from Young Scholar Club in 2005. He received high ranks in Iranian National Electrical Engineering Student Olympiads of 2009 and 2010.

Reza Asvadi (M'11) received the B.Sc. and Ph.D. degrees (Hons.) in electrical engineering from the K. N. Toosi University of Technology, Tehran, Iran, and the M.Sc. degree from the Sharif University of Technology, Tehran. He was a Post-Doctoral Researcher with the University of Oulu, Oulu, Finland. He is an Assistant Professor with the Department of Electrical Engineering, Shahid Beheshti University, Tehran. His research interests include coding theory, information theory, signal processing, and PHY-layer secrecy. He has published more than ten journal articles on these topics in prestigious international journals.

Jun Chen (S'03–M'06–SM'16) received the B.E. degree (Hons.) in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2004 and 2006, respectively.

He was a Post-Doctoral Research Associate with the Coordinated Science Laboratory, University of Illinois at Urbana–Champaign, Urbana, IL, USA, from September 2005 to July 2006, and a Post-Doctoral Fellow with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, from July 2006 to August 2007. Since September 2007, he has been with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is currently a Professor. His research interests include information theory, machine learning, wireless communications, and signal processing.

Dr. Chen held the title of the Barber-Gennum Chair in Information Technology from 2008 to 2013 and the Joseph Ip Distinguished Engineering Fellow from 2016 to 2018. He was a recipient of the Josef Raviv Memorial Postdoctoral Fellowship in 2006, the Early Researcher Award from the Province of Ontario in 2010, and the IBM Faculty Award in 2010. He served as an Associate Editor for IEEE TRANSACTIONS ON INFORMATION THEORY from 2014 to 2016.

Mahmoud Ahmadian-Attari received the M.Eng. degree in electrical engineering from the University of Tehran in 1977 and the Ph.D. degree in electrical engineering from the University of Manchester in 1996.

He joined the K.N.Toosi University of Technology, Tehran, Iran, in 1990, where he is currently a Professor and the Head of the Coding and Cryptography Laboratory (CCL), Department of Electrical Engineering. He has held positions at Teacher Training University (currently Kharazmi University), Alzahra University, and the Ministry of Culture and Higher Education (currently Ministry of Science, Research and Technology). He is the author of the book *Error Control Codes in Telecommunication Systems* in Persian (K. N. Toosi University: Press, 2013). His main professional interests are within the areas of error correcting codes, cryptography, and secure communications. He has published some 70 journal articles on these topics.

Dr. Ahmadian-Attari has been a member of Communications and signal Processing Chapter of IEEE Iran Section, since 2013. He is active in conference organization, most recently as the General Chair of the ISCISC'2010 Conference on Information Security and Cryptology, the Technical Program Committee in IWCIT'2015, IWCIT'2014, IWCIT'2013, ISCISC'2011, ISCISC'2012, ISCISC'2013, ISCISC'2014, ISCISC'2015, ISCISC'2016, ISCISC'2017, ISCISC'2018, and ISCISC'2019.

Tad Matsumoto (S'84–M'98–F'10) received the B.S. and M.S. degrees in electrical engineering under the supervision of Prof. Sin-Ichi Takahashi and the Ph.D. degree in electrical engineering under the supervision of Prof. Masao Nakagawa from Keio University, Yokohama, Japan, in 1978, 1980, and 1991, respectively. He joined Nippon Telegraph and Telephone Corporation (NTT) in 1980, where he was involved in a lot of research and development projects of mobile wireless communications systems. In 1992, he transferred to NTT DoCoMo, where he researched on code-division multiple-access techniques for mobile communication systems. In 1994, he transferred to NTT America, where he served as a Senior Technical Advisor for a joint project between NTT and NEXTEL Communications. In 1996, he returned to NTT DoCoMo, where he served as the Head of the Radio Signal Processing Laboratory, until 2001. He researched on adaptive signal processing, multiple-input multiple-output turbo signal detection, interference cancellation, and space-time coding techniques for broadband mobile communications. In 2002, he moved to the University of Oulu, Finland, where he served as a Professor with the Centre for Wireless Communications. In 2006, he was a Visiting Professor with the Ilmenau University of Technology, Ilmenau, Germany, supported by the German MERCATOR Visiting Professorship Program. Since 2007, he has been a Professor with the Japan Advanced Institute of Science and Technology, Japan, while keeping a cross-appointment position with the University of Oulu (Currently, the cross appointment agreement is frozen). He is a member of the IEICE. He was a recipient of the IEEE VTS Outstanding Service Award in 2001, the Nokia Foundation Visiting Fellow Scholarship Award in 2002, the IEEE Japan Council Award for Distinguished Service to the Society in 2006, the IEEE Vehicular Technology Society James R. Evans Avant Garde Award in 2006, the Thuringen State Research Award for Advanced Applied Science in 2006, the 2007 Best Paper Award of the Institute of Electrical, Communication, and Information Engineers of Japan, in 2008, the Telecom System Technology Award from the Telecommunications Advancement Foundation in 2009, the IEEE COMMUNICATION LETTERS Exemplary Reviewer in 2011, the Nikkei Wireless Japan Award in 2013, the IEEE VTS Recognition for Outstanding Distinguished Lecturer in 2016, and the IEEE TRANSACTIONS ON COMMUNICATIONS Exemplary Reviewer in 2018. He has been serving as an IEEE Vehicular Technology Distinguished Speaker since 2016. He has led a lot of projects supported by the Academy of Finland, European FP7, and the Japan Society for the Promotion of Science and Japanese private companies. He was appointed as a Finland Distinguished Professor from 2008 to 2012, supported by Finnish National Technology Agency (Tekes) and Finnish Academy, under which he preserves the rights to participate in and apply for European and Finnish National Projects.