Two-Way Source-Channel Coding

Jian-Jia Weng, *Member, IEEE*, Fady Alajaji, *Senior Member, IEEE*, and Tamás Linder, *Fellow, IEEE*

***Abstract*— We propose an adaptive lossy joint source-channel coding (JSCC) scheme for sending correlated sources over two-terminal discrete-memoryless two-way channels (DM-TWCs). The main idea is to couple the independent operations of the terminals via an adaptive coding mechanism, which can mitigate cross-interference resulting from simultaneous channel transmissions and concurrently exploit the sources' correlation to reduce the end-to-end reconstruction distortions. Our adaptive JSCC scheme not only subsumes existing lossy coding methods for two-way simultaneous communication but also improves their performance. Furthermore, we derive outer bounds for our two-way lossy transmission problem and establish complete JSCC theorems in some special settings. In these special cases, a non-adaptive separate source-channel coding (SSCC) scheme achieves the optimal performance, thus simplifying the design of the source-channel communication system.**

***Index Terms*— Network information theory, two-way channels, lossy transmission, joint source-channel coding, correlated sources, hybrid analog and digital coding, superposition coding, adaptive coding.**

I. Introduction

SHANNON'S two-way communication [4] considers full-duplex data transmission between two terminals. The terminals can send and receive data simultaneously on a shared two-way channel (TWC) without multiplexing [5] to make the best utilization of channel resources. The TWC was recently used as a building block in the construction of high spectral-efficiency transmission systems [6]–[8]. However, designing an adaptive coding scheme for simultaneous transmission on TWCs remains challenging. More precisely, adaptive coding generates current channel inputs by taking into consideration past received signals. This mechanism conceptually improves the system's performance, but finding optimal coding methods remains elusive.

In this paper, we investigate the adaptive coding problem from an information-theoretic perspective. Specifically, we consider the two-way lossy source-channel communication system depicted in Fig 1, where two terminals exchange correlated sources on a discrete-memoryless TWC (DM-TWC). Our objective is to characterize the achievable distortion region of the system for any given correlated sources, memoryless channel, transmission rate, and distortion measures. Before presenting our contributions, we first review existing results on two-way channel coding and source coding.

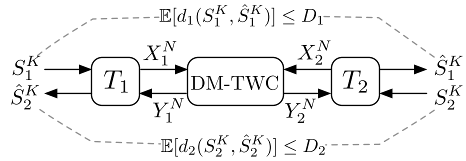


Fig. 1. The block diagram for the lossy transmission of correlated source pair (S_1^K, S_2^K) via N uses of a noisy DM-TWC.

A. Literature Review

The capacity problem for general DM-TWCs is not yet completely solved in single-letter form. In [4], Shannon presented a random coding inner bound and a cut-set outer bound to the capacity region. He also exploited channel symmetry properties [4, Section 11] to determine the capacity region in some special cases, which are further generalized in [9]–[14]. For DM-TWCs with symmetry properties, it was shown that Shannon's inner bound is tight, and hence adaptive coding is not needed to achieve capacity. In the literature, there are other improved inner bounds [15]–[19] and outer bounds [9], [20]. A common idea to improve on Shannon's inner bound is to coordinate the terminals' transmission via a stationary process. Although the terminals operate independently, the adaptive encoding procedure driven by the stationary process ultimately coordinates their encoding operations, thus jointly optimizing their transmissions. In the improved outer bounds, one typically seeks extra dependency among channel inputs.

In two-terminal two-way lossy source coding, the DM-TWC in Fig. 1 is assumed to be noiseless. In [21], Kaspi established a rate-distortion (RD) region for this system,[1] which characterizes the trade-off between source compression rate and distortion, under an interactive communication protocol. Specifically, the protocol divides the entire transmission period into small segments, and only one terminal sends data at each segment. With this protocol, each terminal can decode a coarse description of the other terminal's messages after observing a

Manuscript received June 26, 2020; revised March 22, 2021; accepted July 17, 2021. Date of publication July 27, 2021; date of current version September 15, 2021. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada. This article was presented in part at the 2017 IEEE International Workshop on Information Theory [1], in part at the 2019 IEEE International Symposium on Information Theory [2], and in part at the 2020 IEEE International Symposium on Information Theory [3]. *(Corresponding author: Jian-Jia Weng.)*

Jian-Jia Weng is with the Department of Electrical Engineering, National Taiwan Ocean University, Keelung 202301, Taiwan (e-mail: jianjiaweng@ieee.org).

Fady Alajaji and Tamás Linder are with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: fady@mast.queensu.ca; linder@mast.queensu.ca).

Communicated by C. Tian, Associate Editor for Source Coding.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3100781.

Digital Object Identifier 10.1109/TIT.2021.3100781

[1]Kaspi's original proof relies on tree codes using an intricate approach. A simper proof can be found in [22, Section 20.3.3] based on the Wyzer-Ziv source coding scheme [23].

new segment of channel outputs. All decoded coarse descriptions are then treated as side-information to compress source messages until final reconstructions are obtained. In [24], Maor and Merhav extended Kaspi's result within the application of successive source refinement. Another related two-way source coding problem, where each terminal is only interested in extracting hidden information related to the source messages of the other terminal, is tackled in [25] under the so-called collaborative information bottleneck problem. The rate-relevance trade-off is determined under Kaspi's transmission protocol.

In addition to the above results, there are other extensions of the source coding problem such as two-way source coding with a helper [26], two-way multi-terminal source coding [27], [28], and two-way function computation [29], [30]. The capacity problem was also studied for TWCs with memory [13] and in a multi-terminal setting with more than two terminals such as multi-access/broadcast, Z, and interference TWCs [31] and three-way channels [32]–[34]. These architectures are beyond the scope of this paper.

B. Notation and Problem Setup

We next introduce the notation used in the paper. The symbols $\mathbb{Z}_+$ and $\mathbb{R}_{\geq 0}$ denote the sets of positive integers and non-negative real numbers, respectively. For any $i \geq 1$, let $A^i \triangleq (A_1, A_2, \ldots, A_i)$ denote a length-i sequence of random variables with common alphabet $\mathcal{A}$. The realization of A^i will be denoted by $a^i = (a_1, a_2, \ldots, a_i) \in \mathcal{A}^i$, where $\mathcal{A}^i$ is the i-fold Cartesian product of $\mathcal{A}$. When the length i is clear from the context, we may write $\boldsymbol{A}$ and $\boldsymbol{a}$ in lieu of A^i and a^i, respectively. Throughout the paper, all alphabets are finite, except for the Gaussian case briefly considered in Section VI-A. Moreover, we delineate each terminal by index j or j', where $j, j' \in \{1, 2\}$. To simplify the presentation, we assume that $j \neq j'$ when these indices appear together. Furthermore, the kth source message of terminal j is denoted by $S_{j,k}$, and its reconstruction at terminal j' is given by $\hat{S}_{j,k}$; also, the nth channel input and output of terminal j are denoted by $X_{j,n}$ and $Y_{j,n}$, respectively. For these system variables, we use $\mathcal{S}_j$, $\hat{\mathcal{S}}_j$, $\mathcal{X}_j$, and $\mathcal{Y}_j$ to denote their respective alphabets. The standard notation $\mathbb{E}$ stands for the expectation operator and $\mathbb{1}\{\cdot\}$ stands for the indicator function.

We are now ready to define our problem. As depicted in Fig. 1, two terminals exchange a block of correlated source messages (S_1^K, S_2^K) of length-K via N uses of a noisy TWC. Terminal j only observes S_j^K and intends to reconstruct $S_{j'}^K$ from S_j^K and Y_j^N subject to a distortion constraint. Here, we assume that the source pairs $(S_{1,k}, S_{2,k})$, $1 \leq k \leq K$, are independent and have the common joint probability distribution P_{S_1,S_2}; i.e., $P_{S_1^K,S_2^K}(s_1^K, s_2^K) = \prod_{k=1}^K P_{S_1,S_2}(s_{1,k}, s_{2,k})$, where $(s_{1,k}, s_{2,k}) \in \mathcal{S}_1 \times \mathcal{S}_2$. The distortion for the reconstruction $\hat{s}_j^K$ of source message s_j^K is assessed via $d_j(s_j^K, \hat{s}_j^K) \triangleq K^{-1} \sum_{k=1}^K d_j(s_{j,k}, \hat{s}_{j,k})$, where $d_j : \mathcal{S}_j \times \hat{\mathcal{S}}_j \to \mathbb{R}_{\geq 0}$ is a single-letter distortion measure for source S_j. Furthermore, the noisy TWC is used without adopting any interactive communication protocol such as in [21], [24]. We only consider DM-TWCs with input alphabet $\mathcal{X}_j$ and output alphabet $\mathcal{Y}_j$ for terminal j, $j = 1, 2$, and with transition probability $P_{Y_1,Y_2|X_1,X_2}$. More precisely, we have that $P_{Y_{1,n},Y_{2,n}|X_1^n,X_2^n,Y_1^{n-1},Y_2^{n-1}} = P_{Y_{1,n},Y_{2,n}|X_{1,n},X_{2,n}} = P_{Y_1,Y_2|X_1,X_2}$ for all n. For this system setup, we seek direct (forward) and converse coding theorems for lossy source-channel transmissibility.

C. Related Work and Our Approach

To the best of our knowledge, there are only few works related to our problem setup. In [4, Section 14], Shannon implicitly illustrated that perfect matching among the source and channel statistics and alphabets results in error-free communication, with the optimal scheme given by uncoded transmission. In [24], the JSCC problem was studied for DM-TWCs which consist of two independent one-way channels. Together with the protocol mentioned in Section I-A, Kaspi's source coding result was extended for successive source refinement. Also, a complete JSCC theorem was derived in this particular setting. By contrast, the authors in [35, Section VIII] tackled the two-way transmission problem for general DM-TWCs without deploying any protocol. The correlation-preserving coding scheme of [36] was adopted for almost lossless transmission; i.e., when requiring the block error rate of the source reconstructions to vanish asymptotically. Similar to Shannon's idea, the (non-adaptive) coding scheme of [35] can preserve source correlation in the channel inputs to facilitate two-way transmission; however, it does not apply to the lossy setup. In this paper, we tackle a transmission problem that is more general in many aspects; e.g., we do not consider a particular type of DM-TWC or assume a given communication protocol. We next sketch the concepts behind our main JSCC achievability result.

As the transmissions of the terminals influence each other on a shared channel and generally cause cross-interference, we propose to design the coding strategies jointly.[2] For this purpose, we construct joint source-channel codes to induce a stationary Markov chain that couples all variables of the communication system in Fig. 1. In principle, when the channel inputs are generated by such codes, all system variables will behave according to the stationary distribution of the induced chain, thus coordinating the independent transmissions of the terminals. Specifically, we combine the following coding techniques to build our adaptive codes. First, we adopt the functional form of superposition coding [37] to generate channel inputs, which plays a central role in inducing the desired Markov transmission process. We also modify the analog/digital hybrid coding scheme of [38] to exploit side-information for decoding, in addition to its original source-correlation-preserving mechanism. Moreover, we use past channel inputs and outputs similarly to [15]

[2]As a DM-TWC can be viewed as two interacting one-way channels [13, Section II-C], each terminal's input carries information and selects the transmission (sub-)channel for the other terminal. For example, the information of terminal 1 is transmitted to terminal 2 through the marginal channel $P_{Y_2|X_1,X_2}$ via channel input X_1. However, any specific input $X_1 = x_1$ also concurrently determines the marginal (sub-)channel $P_{Y_1|X_1=x_1,X_2}$ for terminal 2 to transmit information. By mitigating cross-interference, we mean that the encoders of the two terminals generate appropriate correlated channel inputs that simultaneously facilitate the channel selections and the information transmissions on the selected (sub-)channels.

to enable adaptive coding. We note that although these techniques are not new, combining and integrating them into an adaptive two-way coding framework for our problem setup is challenging, especially to ensure stationarity of the system variables. We next summarize the contributions of the paper.

D. Summary of Contributions

Our primary contribution is the construction of an adaptive coding scheme to prove a direct JSCC theorem; but we also derive some converse results and complete JSCC theorems. The details are as follows.

• *Inner Bounds and Examples*: a general JSCC result (Theorem 1) in single-letter form for two-way lossy simultaneous transmission is established using the concepts of superposition coding, hybrid analog/digital coding, and adaptive channel coding, together with a low-complexity sliding-window decoder. Two simplified achievability results (Corollaries 1 and 2) are derived from the main theorem. Moreover, our coding method is shown to subsume some basic schemes such as uncoded transmission and the concatenation of Wyner-Ziv (WZ) source coding and Shannon's (or Han's) channel coding; it also recovers the almost lossless transmission of [35]. Four illustrated examples (Examples 1-4) are provided to highlight the difference between the coding schemes. Overall, not only do we create a unified framework for the prior coding methods but we also improve on their performance.

• *Outer Bounds and Complete JSCC Theorems*: two outer bounds (Lemmas 1 and 2) to the achievable distortion region are obtained using standard arguments. The bounds are expressed in terms of the standard RD function and the conditional RD function and are hence easy to compute for many classical models of correlated sources. Furthermore, four complete theorems (Theorems 2-5) that fully characterize the achievable distortion region for certain system settings are obtained. Specifically, for DM-TWCs with symmetry properties [13], we show the optimality of SSCC in the following settings:

- lossy transmission of independent sources;
- almost lossless transmission of correlated sources;
- lossy transmission of correlated sources whose WZ and conditional RD functions are equal;
- lossy transmission of correlated sources having a common part in the sense of Gács-Körner-Witsenhausen [22, Section 14.2.2].

Examples for Theorems 4 and 5 (Examples 5-7) are also provided.

The rest of the paper is organized as follows. In Section II, definitions and background information are provided. Our direct coding theorem is presented in Section III, and its proof is provided in Appendix A. Simplified versions of the main theorem are given in Section IV, together with a derivation of the associated coding schemes. Section V establishes converse results and complete JSCC theorems. Examples and a discussion are given in Section VI, and conclusions are drawn in Section VII.

II. PRELIMINARIES

In this section, we define joint source-channel codes and the achievable distortion region for source-channel communication over a TWC. We also review various RD function expressions for point-to-point communication and channel coding results for DM-TWCs, which will be used in Section IV.

A. Definitions

For our problem setup, a joint source-channel code is defined as follows.

Definition 1: An (N,K) code for transmitting (S_1^K,S_2^K) over a DM-TWC consists of two sequences of encoding functions $\underline{f}_1 \triangleq \{f_{1,n}\}_{n=1}^N$ and $\underline{f}_2 \triangleq \{f_{2,n}\}_{n=1}^N$ such that

$$f_{1,1}:\mathcal{S}_1^K\to\mathcal{X}_1,\quad f_{1,n}:\mathcal{S}_1^K\times\mathcal{Y}_1^{n-1}\to\mathcal{X}_1$$
$$f_{2,1}:\mathcal{S}_2^K\to\mathcal{X}_2,\quad f_{2,n}:\mathcal{S}_2^K\times\mathcal{Y}_2^{n-1}\to\mathcal{X}_2$$

for $n=2,3,\dots,N$, and two decoding functions $g_1:\mathcal{S}_1^K\times\mathcal{Y}_1^N\to\hat{\mathcal{S}}_2^K$ and $g_2:\mathcal{S}_2^K\times\mathcal{Y}_2^N\to\hat{\mathcal{S}}_1^K$.

The channel inputs at time $n=1$ are only functions of the source messages, i.e., $X_{j,1}=f_{j,1}(S_j^K)$, but the subsequent channel inputs are generated by also adapting to the previous channel outputs via $X_{j,n}=f_{j,n}(S_j^K,Y_j^{n-1})$ for $n=2,3,\dots,N$. Such encoding strategy is called adaptive coding, in contrast to its non-adaptive counterpart where $X_{j,n}=f_{j,n}(S_j^k)$ for all n. We remark that our code definition also involves block-wise decoding; i.e., terminal j reconstructs $S_{j'}^K$ via $\hat{S}_{j'}^K=g_j(S_j^K,Y_j^N)$ after receiving the entire N channel outputs.

Moreover, the rate of the joint source-channel code is given by K/N (source symbols/channel use), and the associated expected distortion is $D_j(K)\triangleq\mathbb{E}[d_j(S_j^K,\hat{S}_j^K)]$, where the expectation is taken with respect to the joint probability distribution

$$P_{S_1^K,S_2^K,X_1^N,X_2^N,Y_1^N,Y_2^N}=P_{S_1^K,S_2^K}\left(\prod_{n=1}^N P_{X_{1,n}|S_1^K,Y_1^{n-1}}\right)\left(\prod_{n=1}^N P_{X_{2,n}|S_2^K,Y_2^{n-1}}\right)\left(\prod_{n=1}^N P_{Y_{1,n},Y_{2,n}|X_{1,n},X_{2,n}}\right),$$

where $P_{Y_{1,n}Y_{2,n}|X_{1,n},X_{2,n}}=P_{Y_1,Y_2|X_1,X_2}$ for $n=1,2,\dots,N$ (determined by the DM-TWC).

Definition 2: A distortion pair (D_1,D_2) is said to be achievable at rate R if there exists a sequence of (N,K) joint source-channel codes (where N is a function of K) such that $\lim_{K\to\infty}K/N=R$ and $\limsup_{K\to\infty}D_j(K)\le D_j$, $j=1,2$. The achievable distortion region of a rate-R two-way lossy transmission system is the convex closure of the set of all achievable distortion pairs (at rate R).

B. Rate-Distortion Functions

As a DM-TWC can be viewed as two state-dependent one-way channels, the following source coding related functions (each expressed in terms of a constrained minimization of a mutual information quantity) for one-way systems are also useful in the two-way channel setup.

- Standard RD function [22, Sec. 3.6]:

$$R^{(j)}(D_j) = \min_{P_{\hat{S}_j|S_j}:\mathbb{E}[d_j(S_j,\hat{S}_j)]\le D_j} I(S_j;\hat{S}_j). \quad (1)$$

- WZ-RD function [23]: Letting $T_j \in \mathcal{T}_j$ with $|\mathcal{T}_j| \le |\mathcal{S}_j|+1$ denote an auxiliary random variable that satisfies the Markov chain $T_j \multimap S_j \multimap S_{j'}$, we have

$$R^{(j)}_{\mathrm{WZ}}(D_j) = \min_{P_{T_j|S_j}} \min_{\substack{h:\mathcal{T}_j\times\mathcal{S}_{j'}\to\hat{\mathcal{S}}_j \\ \mathbb{E}[d_j(S_j,h(T_j,S_{j'}))]\le D_j}} I(S_j;T_j|S_{j'}). \quad (2)$$

- Conditional RD function [39]:

$$R_{S_j|S_{j'}}(D_j) = \min_{\substack{P_{\hat{S}_j|S_1,S_2} \\ \mathbb{E}[d_j(S_j,\hat{S}_j)]\le D_j}} I(S_j;\hat{S}_j|S_{j'}). \quad (3)$$

We remark that the source coding schemes that achieve the standard RD and WZ-RD functions can be the building blocks of an SSCC scheme for our overall system. For example, terminal j can apply the WZ coding scheme to compress source S_j^K given side-information $S_{j'}^K$. Although the coding scheme that achieves the conditional RD function cannot be applied in our problem setup (since there is no common side-information at the encoder and the decoder in general), the scheme is useful when S_1 and S_2 have a common part in the sense of Gács-Körner-Witsenhausen [22, Section 14.2.2]. We will use this result in Theorem 5 (see Section V-B).

C. Capacity Bounds for DM-TWCs

To introduce capacity bounds for DM-TWCs, we first give some definitions. Roughly speaking, an $(N, R_{\mathrm{c},1}, R_{\mathrm{c},2})$ channel code for a DM-TWC is defined similarly to an (N, K) joint source-channel code, except that the correlated sources S_1^K and S_2^K are replaced with independent and uniformly distributed random indices $I_1 \in \mathcal{I}_1$ and $I_2 \in \mathcal{I}_2$, respectively, where $|\mathcal{I}_1| = 2^{NR_{\mathrm{c},1}}$ and $|\mathcal{I}_2| = 2^{NR_{\mathrm{c},2}}$. As a result, two-way channel codes can incorporate or exclude adaptive coding. Given a DM-TWC, a channel coding rate pair $(R_{\mathrm{c},1}, R_{\mathrm{c},2})$ is called achievable if there exists a sequence of $(N, R_{\mathrm{c},1}, R_{\mathrm{c},2})$ channel codes such that I_1 and I_2 can be reliably exchanged (i.e., with asymptotically vanishing decoding error probability). The capacity region is defined as the convex closure of the set of all achievable rate pairs.

To date, a single-letter characterization of the capacity region of general DM-TWCs is still not found. In [4], Shannon derived the inner bound region

$$\overline{\mathrm{conv}}\left(\bigcup_{P_{X_1}P_{X_2}}\left\{(R_{\mathrm{c},1},R_{\mathrm{c},2}) : R_{\mathrm{c},1} < I(X_1;Y_2|X_2),\ R_{\mathrm{c},2} < I(X_2;Y_1|X_1)\right\}\right) \quad (4)$$

and the outer bound region

$$\bigcup_{P_{X_1,X_2}}\left\{(R_{\mathrm{c},1},R_{\mathrm{c},2}) : R_{\mathrm{c},1} < I(X_1;Y_2|X_2),\ R_{\mathrm{c},2} < I(X_2;Y_1|X_1)\right\}$$

for channel capacity, where $\overline{\mathrm{conv}}(\cdot)$ denotes taking the closure of the convex hull. In general, the two capacity bounds do not coincide, but they match each other for channels with symmetry properties; i.e., DM-TWCs that satisfy the channel symmetry conditions in either [13, Theorem 1] or [13, Theorem 4]. For these "symmetric" DM-TWCs, the capacity region can be exactly determined via non-adaptive coding and is given by the set of all achievable rate pairs in (4) under independent inputs. Moreover, taking the convex closure in (4) is not needed.

Shannon's inner bound result was later improved by Han [15] under an adaptive channel coding scheme, showing that any rate pair in the following region is achievable:

$$\overline{\mathrm{conv}}\left(\bigcup_{P_{\tilde{V}_1,\tilde{V}_2,\tilde{W}_1,\tilde{W}_2,X_1,X_2}}\left\{(R_{\mathrm{c},1},R_{\mathrm{c},2}) : R_{\mathrm{c},1} < I(\tilde{V}_1;X_2,Y_2,\tilde{V}_2,\tilde{W}_2),\ R_{\mathrm{c},2} < I(\tilde{V}_2;X_1,Y_1,\tilde{V}_1,\tilde{W}_1)\right\}\right)$$

where the joint probability distribution $P_{\tilde{V}_1,\tilde{V}_2,\tilde{W}_1,\tilde{W}_2,X_1,X_2}$ is defined in [15, Section IV].[3] We note that Kramer further generalized Han's result from a concatenated coding perspective [18, Section 4.3.2] with achievable rate pairs obtained in terms of conditional directed mutual information quantities using a random coding error exponent analysis under maximum-likelihood decoding [40]. In this paper, as we pursue single-letter expressions, we mainly focus on Shannon's and Han's results.

III. Direct JSCC Theorem Based on Adaptive Coding

This section establishes the most general achievability result in the paper. Without loss of generality, we only consider rate-one transmission, i.e., $N = K$; other rates can be obtained via suitable super-symbols.[4] The main idea is to employ a stationary Markov chain to coordinate the independent encoding operations of the two terminals. More specifically, we first construct a coded channel in Section III-A to represent the use of the DM-TWC multiple times for adaptive coding purposes; the coded channel also integrates two encoding functions so that the raw inputs of the coded channel are naturally converted into adaptive channel inputs to the original channel. To coordinate the independent transmissions of the two terminals, we next couple the variables of the coded channel at different time instants via a stationary time-homogeneous Markov chain (Sections III-B and III-C). Roughly speaking, when the raw inputs of the coded channel are drawn from the stationary Markov chain, the adaptive channel inputs to the original DM-TWC become coupled. Here we remark that stationarity is key in obtaining an achievable region in single-letter form as given in Theorem 1 in Section III-D. In the following, we formally describe the above technical ingredients used in obtaining Theorem 1.

[3]The random variables $\tilde{V}_j$ and $\tilde{W}_j$ correspond to the random variables $\tilde{U}_j$ and $\tilde{W}_j$ in Han's scheme, respectively.

[4]To obtain a rate-$\frac{K_1}{N_1}$ result, we define a super source symbol (resp., a super channel input/output symbol) by combining K_1 source symbols (resp., N_1 channel input/output symbols).

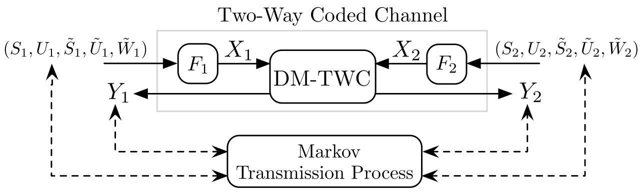


Fig. 2. An illustration of the transmission over the two-way coded channel.

A. Two-Way Coded Channel

Consider an auxiliary coded channel built on the original (physical) DM-TWC, as shown in the central box of Fig. 2. The coded channel has inputs $S_j, U_j, \tilde{S}_j, \tilde{U}_j$ and $\tilde{W}_j$ at terminal j. The input pairs (S_j, U_j) and $(\tilde{S}_j, \tilde{U}_j)$ are used to carry the current and some prior source information, respectively, where U_j (resp., $\tilde{U}_j$) denotes the coded version of S_j (resp., $\tilde{S}_j$). The input $\tilde{W}_j$ carries some past channel inputs and outputs at terminal j. The new channel also involves two encoding functions $F_j : \mathcal{S}_j \times \mathcal{U}_j \times \tilde{\mathcal{S}}_j \times \tilde{\mathcal{U}}_j \times \tilde{\mathcal{W}}_j \to \mathcal{X}_j$, which transform the inputs of the coded channel into the inputs for the original DM-TWC. The outputs of the new channel are still Y_1 and Y_2. The joint input probability distribution of the coded channel is given by

$$\begin{aligned} &P_{S_1,S_2,U_1,U_2,\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2} \\ &\quad = P_{S_1,S_2} P_{U_1|S_1} P_{U_2|S_2} P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2}, \end{aligned} \tag{5}$$

and the transition probability of the coded channel is given by

$$\begin{aligned} &P_{Y_1,Y_2|S_1,S_2,U_1,U_2,\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2} \\ &\qquad (y_1, y_2 | s_1, s_2, u_1, u_2, \tilde{s}_1, \tilde{s}_2, \tilde{u}_1, \tilde{u}_2, \tilde{w}_1, \tilde{w}_2) \\ &= \sum_{x_1,x_2} \mathbb{1}\{x_1 = F_1(s_1, u_1, \tilde{s}_1, \tilde{u}_1, \tilde{w}_1)\} \\ &\quad \cdot \mathbb{1}\{x_2 = F_2(s_2, u_2, \tilde{s}_2, \tilde{u}_2, \tilde{w}_2)\} P_{Y_1,Y_2|X_1,X_2}(y_1, y_2 | x_1, x_2). \end{aligned} \tag{6}$$

B. Markov Chain for the Coded Channel

For the repeated use over time of the two-way coded channel, we next construct a discrete-time Markov chain for the overall system with state space:

$$\begin{aligned} &\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \tilde{\mathcal{S}}_1 \times \tilde{\mathcal{S}}_2 \times \tilde{\mathcal{U}}_1 \times \tilde{\mathcal{U}}_2 \\ &\qquad \times \tilde{\mathcal{W}}_1 \times \tilde{\mathcal{W}}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}_1 \times \mathcal{Y}_2, \end{aligned} \tag{7}$$

where $\tilde{\mathcal{S}}_j \triangleq \mathcal{S}_j, \tilde{\mathcal{U}}_j \triangleq \mathcal{U}_j$, and $\tilde{\mathcal{W}}_j \triangleq \mathcal{X}_j \times \mathcal{Y}_j$ for $j = 1, 2$. This Markov chain will be used to coordinate the transmissions of the two terminals as shown in Fig. 2. Let

$$\begin{aligned} Z^{(t)} \triangleq (&S_1^{(t)}, S_2^{(t)}, U_1^{(t)}, U_2^{(t)}, \tilde{S}_1^{(t)}, \tilde{S}_2^{(t)}, \tilde{U}_1^{(t)}, \tilde{U}_2^{(t)}, \\ &\tilde{W}_1^{(t)}, \tilde{W}_2^{(t)}, X_1^{(t)}, X_2^{(t)}, Y_1^{(t)}, Y_2^{(t)}) \end{aligned} \tag{8}$$

denote the state of the Markov chain at time $t \in \mathbb{Z}_+$, where we set $\tilde{S}_j^{(t)} \triangleq S_j^{(t-1)}$, $\tilde{U}_j^{(t)} \triangleq U_j^{(t-1)}$, and $\tilde{W}_j^{(t)} \triangleq (X_j^{(t-1)}, Y_j^{(t-1)})$. Given a parameter tuple $(P_{U_1|S_1}, P_{U_2|S_2}, P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2}, F_1, F_2)$, we generate the quadruple $(S_1^{(t)}, S_2^{(t)}, U_1^{(t)}, U_2^{(t)})$ for all t according to $P_{S_1,S_2,U_1,U_2} = P_{S_1,S_2} P_{U_1|S_1} P_{U_2|S_2}$ independently of $(\tilde{S}_1^{(t)}, \tilde{S}_2^{(t)}, \tilde{U}_1^{(t)}, \tilde{U}_2^{(t)}, \tilde{W}_1^{(t)}, \tilde{W}_2^{(t)})$. The tuple $(\tilde{S}_1^{(1)}, \tilde{S}_2^{(1)}, \tilde{U}_1^{(1)}, \tilde{U}_2^{(1)}, \tilde{W}_1^{(1)}, \tilde{W}_2^{(1)})$ is initialized according to $P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2}$. The physical channel input at terminal j is naturally produced as $X_j^{(t)} = F_j(S_j^{(t)}, U_j^{(t)}, \tilde{S}_j^{(t)}, \tilde{U}_j^{(t)}, \tilde{W}_j^{(t)})$, and the received channel output is $Y_j^{(t)}$. Based on this construction, the transition kernel of $\{Z^{(t)}\}$ is given in (9), as shown at the bottom of the page, for $t \geq 2$. It is easy to see that the process $\{Z^{(t)}\}$ is a first-order time-homogeneous Markov chain. However, whether or not the chain is stationary depends on the given parameters.

C. Stationary Distribution Under Distortion Constraints

To obtain an achievability result with time-independent conditions, we only consider a stationary Markov chain. The following procedure can be used to find its parameters. Given P_{S_1,S_2} and $P_{Y_1,Y_2|X_1,X_2}$, we first fix a choice of $P_{U_j|S_j}$ and F_j, $j = 1, 2$, and write the transition kernel (9) in matrix form as Q_Z. The matrix Q_Z is stochastic, and since all alphabets are finite, an eigenvector of Q_Z associated with the eigenvalue 1 exists and gives a stationary distribution P_Z for $\{Z^{(t)}\}$, i.e., $P_Z = P_Z Q_Z$. Clearly, using the marginal distribution $P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2}$ of P_Z with the chosen $P_{U_j|S_j}$ and F_j, $j = 1, 2$, to initialize the Markov chain ensures stationarity. Note that for the stationary chain the two independent quadruples $(S_1^{(t)}, S_2^{(t)}, U_1^{(t)}, U_2^{(t)})$ and $(\tilde{S}_1^{(t)}, \tilde{S}_2^{(t)}, \tilde{U}_1^{(t)}, \tilde{U}_2^{(t)})$ have identical distributions for all t; thus $P_{S_1,S_2,U_1,U_2} = P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}$. Moreover, due to our construction of $\{Z^{(t)}\}$, we have the following necessary conditions for stationarity

$$P_{S_1,S_2} = P_{\tilde{S}_1,\tilde{S}_2}, \tag{10}$$

$$P_{U_j|S_j} = P_{\tilde{U}_j|\tilde{S}_j}, \tag{11}$$

for $j = 1, 2$. For source reconstruction, we next associate the parameters with decoding functions[5] $G_j : \tilde{\mathcal{U}}_{j'} \times \mathcal{S}_j \times \mathcal{U}_j \times \tilde{\mathcal{S}}_j \times \tilde{\mathcal{U}}_j \times \tilde{\mathcal{W}}_j \times \mathcal{Y}_j \to \hat{\tilde{\mathcal{S}}}_{j'}$, $j = 1, 2$. For simplicity, we call the tuple $(P_{U_1|S_1}, P_{U_2|S_2}, P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}, P_{\tilde{W}_1,\tilde{W}_2|\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}, F_1, F_2, G_1, G_2)$ a *configuration*, which specifies a stationary distribution P_Z with the following fac-

[5]As will be seen at the end of the section or in Appendix A, terminal j reconstructs the prior source message $\tilde{S}_{j'}$ as $\hat{\tilde{S}}_{j'}$ after recovering $\tilde{U}_{j'}$; this reconstruction is done via G_j.

$$\begin{aligned} &P_{Z^{(t)}|Z^{(t-1)}}(s_1, s_2, u_1, u_2, \tilde{s}_1, \tilde{s}_2, \tilde{u}_1, \tilde{u}_2, \tilde{w}_1, \tilde{w}_2, x_1, x_2, y_1, y_2 | s_1', s_2', u_1', u_2', \tilde{s}_1', \tilde{s}_2', \tilde{u}_1', \tilde{u}_2', \tilde{w}_1', \tilde{w}_2', x_1', x_2', y_1', y_2') \\ &\quad = P_{S_1,S_2}(s_1, s_2) P_{U_1|S_1}(u_1|s_1) P_{U_2|S_2}(u_2|s_2) \mathbb{1}\{\tilde{s}_1 = s_1'\} \mathbb{1}\{\tilde{s}_2 = s_2'\} \mathbb{1}\{\tilde{u}_1 = u_1'\} \mathbb{1}\{\tilde{u}_2 = u_2'\} \mathbb{1}\{\tilde{w}_1 = (x_1', y_1')\} \\ &\qquad \cdot \mathbb{1}\{\tilde{w}_2 = (x_2', y_2')\} \mathbb{1}\{x_1 = F_1(s_1, u_1, \tilde{s}_1, \tilde{u}_1, \tilde{w}_1)\} \mathbb{1}\{x_2 = F_2(s_2, u_2, \tilde{s}_2, \tilde{u}_2, \tilde{w}_2)\} P_{Y_1,Y_2|X_1,X_2}(y_1, y_2 | x_1, x_2) \end{aligned} \tag{9}$$

torization

$$\underbrace{P_{S_1,S_2}P_{U_1|S_1}P_{U_2|S_2}}_{=P_{S_1,S_2,U_1,U_2}}\underbrace{P_{\tilde{S}_1,\tilde{S}_2}P_{\tilde{U}_1|\tilde{S}_1}P_{\tilde{U}_2|\tilde{S}_2}}_{=P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}}P_{\tilde{W}_1,\tilde{W}_2|\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}$$
$$\cdot P_{X_1|S_1,U_1,\tilde{S}_1,\tilde{U}_1,\tilde{W}_1}P_{X_2|S_2,U_2,\tilde{S}_2,\tilde{U}_2,\tilde{W}_2}P_{Y_1,Y_2|X_1,X_2},$$

where P_{S_1,S_2} and $P_{Y_1,Y_2|X_1,X_2}$ are fixed by the problem setup and $P_{X_j|S_j,U_j,\tilde{S}_j,\tilde{U}_j,\tilde{W}_j}$ is determined by F_j, $j=1,2$. We also let $\Pi_Z(D_1,D_2)$ denote the set of all configurations that induce a stationary chain and satisfy the distortion constraints: $\mathbb{E}[d_j(\tilde{S}_j,\hat{\tilde{S}}_j)] \leq D_j$ for $j=1,2$.

D. Main Result: JSCC Achievability

Based on the above setup, we establish the achievability result in Theorem 1 below; the proof is given in Appendix A. In Theorem 1, one can further convexify the achievable distortion region via a standard time-sharing argument [42].

Theorem 1 (Adaptive JSCC): A distortion pair (D_1,D_2) is achievable for the rate-one lossy transmission of correlated sources over a DM-TWC if there exists a configuration in $\Pi_Z(D_1,D_2)$ such that

$$I(\tilde{S}_1;\tilde{U}_1) < I(\tilde{U}_1;S_2,U_2,\tilde{S}_2,\tilde{U}_2,\tilde{W}_2,X_2,Y_2), \quad (12a)$$
$$I(\tilde{S}_2;\tilde{U}_2) < I(\tilde{U}_2;S_1,U_1,\tilde{S}_1,\tilde{U}_1,\tilde{W}_1,X_1,Y_1). \quad (12b)$$

To facilitate the understanding of the conditions in (12a), we sketch our coding method used in the proof, which extends the hybrid analog/digital coding scheme of [38], used in conjunction with superposition block Markov encoding [15], [41] and a sliding-window decoder, as shown in Fig. 3. In our method, instead of exchanging a single block of source message pairs (S_1^K,S_2^K) via K channel uses, we exchange B blocks of such source message pairs via $K(B+1)$ channel uses for some $B\in\mathbb{Z}_+$. The overall transmission rate is $\frac{B}{B+1}$, which approaches 1 as B goes to infinity. The extra K channel uses can be viewed as added redundancy for data protection.

For $1\leq b\leq B$, let $\boldsymbol{S}_j^{(b)}=(S_{j,1}^{(b)},S_{j,2}^{(b)},\ldots,S_{j,K}^{(b)})$ denote the bth source message block at terminal j; the same indexing convention applies to other variables. As shown in Fig. 3(a),[6] the encoding involves hybrid analog/digital coding, superposition coding, and adaptive channel coding. In the bth transmission block, terminal j first encodes its source message $\boldsymbol{S}_j^{(b)}$ into the digital codeword $\boldsymbol{U}_j^{(b)}$. Then, the current information $(\boldsymbol{S}_j^{(b)},\boldsymbol{U}_j^{(b)})$ and the prior information $(\boldsymbol{S}_j^{(b-1)},\boldsymbol{U}_j^{(b-1)})$ and $(\boldsymbol{X}_j^{(b-1)},\boldsymbol{Y}_j^{(b-1)})$ are combined to generate the channel input $\boldsymbol{X}_j^{(b)}$.

To reconstruct source messages, we employ a sliding-window decoder as depicted in Fig. 3(b). The decoder is designed to operate on two consecutive transmission blocks, but each time it only decodes the earlier source block. For $2\leq b\leq B+1$, suppose that the decoding window is now across the $(b-1)$st and the bth transmission blocks. Given that terminal j has successfully recovered $\boldsymbol{U}_{j'}^{(b')}$ and reconstructed $\boldsymbol{S}_{j'}^{(b')}$ for all $b'<b-1$, the decoder uses all available information in the $(b-1)$st and the bth blocks to recover $\boldsymbol{U}_{j'}^{(b-1)}$ and reconstructs $\boldsymbol{S}_{j'}^{(b-1)}$ as $\hat{\boldsymbol{S}}_{j'}^{(b-1)}$ via G_j. Then, the decoder moves to the bth and the $(b+1)$st blocks to reconstruct $\boldsymbol{S}_{j'}^{(b)}$.

With the above sketch, the left-hand-side and the right-hand-side of (12a) can be interpreted as source compression rates and as transmission rates for reliable communication, respectively. Moreover, the appearance of $(\tilde{S}_j,\tilde{U}_j)$ (rather than (S_j,U_j)) on the left-hand-side of (12a) is due to the sliding-window decoder. The tuple $(S_j,U_j,\tilde{S}_j,\tilde{U}_j,\tilde{W}_j,X_j,Y_j)$ on the right-hand-side of (12a) also illuminates the fact that the decoder at terminal j uses all information within two blocks to decode $\tilde{U}_{j'}$. The detailed coding scheme and the formal proof is provided in Appendix A.

In ending this section, we remark that the unified random coding theorem by Lee and Chung in [43] can be employed to derive achievability results for our source-channel communication based on an unfolded network perspective. Although this approach is powerful due to its generality, the main theorem [43, Theorem 1] is not guaranteed to provide a single-letter characterization of the achievable region. Here, the main challenge lies in finding appropriate coding parameters to induce a desired stationarity property for all involved random variables. In network information theory, one often encounters a trade-off between the tightness of an achievable region and its characterization complexity. Our work aims to improve the existing single-letter results. A by-product is a demonstration of choosing coding parameters for the unified random coding theorem of [43] to obtain achievability conditions in single-letter form. In the next section, we simplify the expressions in (12a) by imposing some encoding constraints. Examples illustrating Theorem 1 will be given in Section VI.

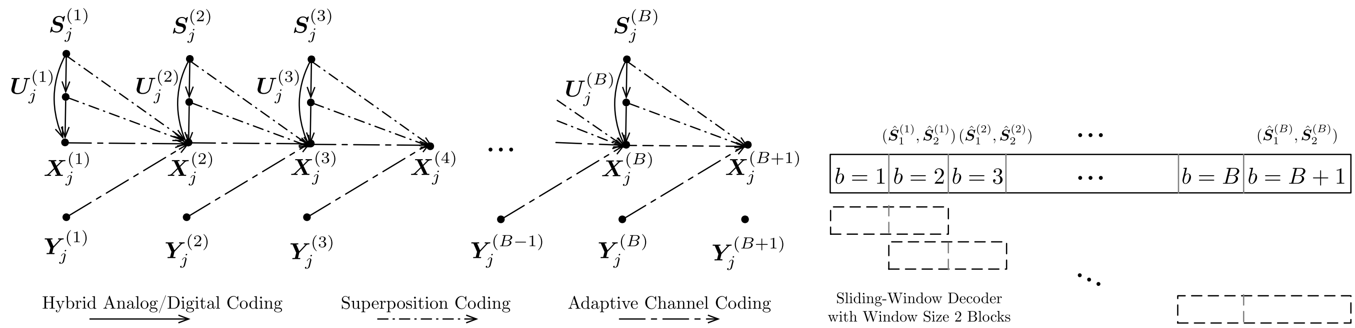
IV. Simplified Configurations and Special Cases

In this section, we consider two simplified forms of encoding to derive special cases from Theorem 1. Our objective is not only to obtain simpler achievability conditions but also to recover existing direct coding theorems for our problem setup. By-products of the derivation are reduced-complexity coding schemes in those special cases. As we will see later in Section V-B, the reduced-complexity schemes in the special cases are sometimes optimal in the sense that the associated achievable distortion region matches a certain outer bound; i.e., the scheme provides a complete JSCC theorem. In such a case, optimal performance can be achieved by a less complex coding scheme. To ease our presentation, we will not refer to the probability distributions P_{S_1,S_2} and $P_{Y_1,Y_2|X_1,X_2}$ in the following result statements as they are fixed and given by the problem setup. Also, we continue to focus on the rate-one case.

A. A Non-Adaptive JSCC Scheme

Our first simplification disables the superposition and adaptive coding components, i.e., we let $X_j=F_j(S_j,U_j,\tilde{S}_j,$

[6]To simplify the presentation of our encoding scheme, we write $\boldsymbol{S}_j^{(b-1)}$, $\boldsymbol{U}_j^{(b-1)}$, and $(\boldsymbol{X}_j^{(b-1)},\boldsymbol{Y}_j^{(b-1)})$ in lieu of $\tilde{\boldsymbol{S}}_j^{(b)}$, $\tilde{\boldsymbol{U}}_j^{(b)}$, and $\tilde{\boldsymbol{W}}_j^{(b)}$, respectively, to refer to the prior information variables at block instance b, for $2\leq b\leq B+1$. Later, when presenting our decoder, we also use $\hat{\boldsymbol{S}}_j^{(b-1)}$ (resp., $\hat{\boldsymbol{U}}_j^{(b-1)}$) rather than $\hat{\tilde{\boldsymbol{S}}}_j^{(b)}$ (resp., $\hat{\tilde{\boldsymbol{U}}}_j^{(b)}$) to denote the reconstruction of $\tilde{\boldsymbol{S}}_j^{(b)}$ (resp., $\tilde{\boldsymbol{U}}_j^{(b)}$).



(a) The encoding process of terminal j, where each node represents a block of variables and each node is a function of other nodes specified by the incoming edges.

(b) The block diagram for sliding-window decoding.

Fig. 3. An illustration of the proposed JSCC method.

$\tilde{U}_j, \tilde{W}_j) \triangleq f_j(\tilde{S}_j, \tilde{U}_j)$ and $\hat{\tilde{S}}_{j'} = G_j(\tilde{U}_{j'}, S_j, U_j, \tilde{S}_j, \tilde{U}_j, \tilde{W}_j, Y_j) \triangleq g_j(\tilde{U}_{j'}, \tilde{S}_j, \tilde{U}_j, Y_j)$ for some f_j and g_j, $j = 1, 2$. Set $P_{\tilde{S}_1, \tilde{S}_2} = P_{S_1, S_2}$, and set $P_{\tilde{U}_j|\tilde{S}_j} = P_{U_j|S_j}$ for a chosen $P_{U_j|S_j}$, $j = 1, 2$, so that (10) and (11) holds. We also set the pair $(\tilde{W}_1, \tilde{W}_2)$ to be *independent* of $(S_1, S_2, U_1, U_2, \tilde{S}_1, \tilde{S}_2, \tilde{U}_1, \tilde{U}_2, X_1, X_2, Y_1, Y_2)$ with joint probability distribution given by

$$
\begin{aligned}
P_{\tilde{W}_1, \tilde{W}_2}(\tilde{w}_1, \tilde{w}_2) = & \sum_{a_1 \in \mathcal{S}_1, a_2 \in \mathcal{S}_2, b_1 \in \mathcal{U}_1, b_2 \in \mathcal{U}_2} P_{\tilde{S}_1, \tilde{S}_2}(a_1, a_2) \\
& \cdot P_{\tilde{U}_1|\tilde{S}_1}(b_1|a_1) P_{\tilde{U}_2|\tilde{S}_2}(b_2|a_2) \mathbb{1}\{\tilde{x}_1 = f_1(a_1, b_1)\} \\
& \cdot \mathbb{1}\{\tilde{x}_2 = f_2(a_2, b_2)\} P_{Y_1, Y_2|X_1, X_2}(\tilde{y}_1, \tilde{y}_2|\tilde{x}_1, \tilde{x}_2). \quad (13)
\end{aligned}
$$

With the above setting, one can directly verify that

$$
\begin{aligned}
P_Z = & P_{S_1, S_2} P_{U_1|S_1} P_{U_2|S_2} P_{\tilde{S}_1, \tilde{S}_2} P_{\tilde{U}_1|\tilde{S}_1} P_{\tilde{U}_2|\tilde{S}_2} P_{\tilde{W}_1, \tilde{W}_2} \\
& P_{X_1|\tilde{S}_1, \tilde{U}_1} P_{X_2|\tilde{S}_2, \tilde{U}_2} P_{Y_1, Y_2|X_1, X_2} \quad (14)
\end{aligned}
$$

is a stationary distribution, i.e., $P_Z = Q_Z P_Z$. Given such P_Z, suppose that the chosen g_j attains distortion level D_j, $j = 1, 2$, so that the configuration $(P_{U_1|S_1}, P_{U_2|S_2}, P_{\tilde{S}_1, \tilde{S}_2, \tilde{U}_1, \tilde{U}_2}, P_{\tilde{W}_1, \tilde{W}_2}, f_1, f_2, g_1, g_2)$ is in $\Pi_Z(D_1, D_2)$. For simplicity, we define the set $\Pi'_Z(D_1, D_2) \subset \Pi_Z(D_1, D_2)$ as the one that contains all such special configurations. Using $\Pi'_Z(D_1, D_2)$, Theorem 1 reduces to the following corollary.

Corollary 1 (Non-Adaptive Hybrid Coding): A distortion pair (D_1, D_2) is achievable for the rate-one lossy transmission of correlated sources over a DM-TWC if there exists a configuration in $\Pi'_Z(D_1, D_2)$ such that

$$
I(\tilde{S}_1; \tilde{U}_1|\tilde{S}_2, \tilde{U}_2) < I(\tilde{U}_1; Y_2|\tilde{S}_2, \tilde{U}_2), \quad (15a)
$$

$$
I(\tilde{S}_2; \tilde{U}_2|\tilde{S}_1, \tilde{U}_1) < I(\tilde{U}_2; Y_1|\tilde{S}_1, \tilde{U}_1). \quad (15b)
$$

Proof: Since $\tilde{U}_{j'}$ is independent of (S_j, U_j) and by definition $\tilde{W}_j$ is independent of $(\tilde{S}_{j'}, S_j, U_j, \tilde{S}_j, \tilde{U}_j, X_j, Y_j)$ for $j = 1, 2$, we can remove $(S_j, U_j, \tilde{W}_j)$ from (12a) without changing the values on the right-hand-side of (12a), e.g.,

$$
\begin{aligned}
& I(\tilde{U}_1; S_2, U_2, \tilde{S}_2, \tilde{U}_2, \tilde{W}_2, X_2, Y_2) \\
& = I(\tilde{U}_1; \tilde{S}_2, \tilde{U}_2, X_2, Y_2) + \underbrace{I(\tilde{U}_1; S_2, U_2, \tilde{W}_2|\tilde{S}_2, \tilde{U}_2, X_2, Y_2)}_{=0}.
\end{aligned}
$$

For (12a), we then have the derivation at the bottom of this page, where the two equalities in (16), shown at the bottom of the page, hold since $X_2 = f_2(\tilde{S}_2, \tilde{U}_2)$. By symmetry, one can analogously deduce (15b) from (12b). ■

We remark that Corollary 1 further subsumes several special cases. In the following derivations, we will show that our chosen parameters form a configuration in $\Pi'_Z(D_1, D_2)$. As $P_{\tilde{W}_1, \tilde{W}_2}$ can be determined via (13) given other parameters, we will not specify $P_{\tilde{W}_1, \tilde{W}_2}$ for the sake of simplicity.

(i) *Uncoded Transmission Scheme*: Strictly speaking, the achievability result of an uncoded scheme cannot be deduced from Corollary 1 since the conditions in (15) have no impact on the scheme's performance. Nevertheless, we still can view it as a special case since every uncoded scheme can be converted into a configuration in our setup, which implies that our coding scheme (used to prove Theorem 1) can emulate uncoded transmission and attains the same distortion levels. Specifically, let $\mathcal{X}_j = \mathcal{S}_j$, $j = 1, 2$. Given encoding functions $\tilde{f}_j$ and decoding functions $\tilde{g}_j$ of an uncoded scheme such that $\mathbb{E}[d_j(\tilde{S}_j, \hat{\tilde{S}}_j)] \le D_j$, we set $X_j = f_j(\tilde{U}_j, \tilde{S}_j) = \tilde{f}_j(\tilde{S}_j)$ and $\hat{\tilde{S}}_j = g_{j'}(\tilde{U}_j, \tilde{S}_{j'}, \tilde{U}_{j'}, Y_{j'}) = \tilde{g}_{j'}(\tilde{S}_{j'}, Y_{j'})$. Also, set $P_{\tilde{S}_1, \tilde{S}_1} = P_{S_1, S_2}$ and

$$
\begin{aligned}
& & I(\tilde{S}_1; \tilde{U}_1) &< I(\tilde{U}_1; \tilde{S}_2, \tilde{U}_2, X_2, Y_2) \\
\Leftrightarrow & & H(\tilde{U}_1) - H(\tilde{U}_1|\tilde{S}_1) &< I(\tilde{U}_1; \tilde{S}_2, \tilde{U}_2) + I(\tilde{U}_1; X_2, Y_2|\tilde{S}_2, \tilde{U}_2) \\
\Leftrightarrow & & H(\tilde{U}_1) - H(\tilde{U}_1|\tilde{S}_1, \tilde{S}_2, \tilde{U}_2) &< H(\tilde{U}_1) - H(\tilde{U}_1|\tilde{S}_2, \tilde{U}_2) + I(\tilde{U}_1; X_2, Y_2|\tilde{S}_2, \tilde{U}_2) \\
\Leftrightarrow & & H(\tilde{U}_1|\tilde{S}_2, \tilde{U}_2) - H(\tilde{U}_1|\tilde{S}_1, \tilde{S}_2, \tilde{U}_2) &< \underbrace{I(\tilde{U}_1; X_2|\tilde{S}_2, \tilde{U}_2)}_{=0} + \underbrace{I(\tilde{U}_1; Y_2|X_2, \tilde{S}_2, \tilde{U}_2)}_{=I(\tilde{U}_1; Y_2|\tilde{S}_2, \tilde{U}_2)} \quad (16) \\
\Leftrightarrow & & I(\tilde{S}_1; \tilde{U}_1|\tilde{S}_2, \tilde{U}_2) &< I(\tilde{U}_1; Y_2|\tilde{S}_2, \tilde{U}_2),
\end{aligned}
$$

$U_j = \tilde{U}_j =$ constant. This setting determines $P_{U_j|S_j}$ and $P_{\tilde{U}_j|\tilde{S}_j}$ uniquely and satisfies (10) and (11). We further obtain $P_{\tilde{W}_1,\tilde{W}_2}$ via (13). Clearly, the configuration $(P_{U_1|S_1}, P_{U_2|S_2}, P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}, P_{\tilde{W}_1,\tilde{W}_2}, \tilde{f}_1, \tilde{f}_2, \tilde{g}_1, \tilde{g}_2)$ belongs to $\Pi'_Z(D_1, D_2)$. Thus, one can establish the achievability result of uncoded transmission in our setup by giving appropriate functions $\tilde{f}_j$ and $\tilde{g}_j$. A more detailed performance analysis for this scheme can be found in [1].

(ii) *SSCC for the Lossy Transmission of Independent Sources*: To satisfy (10), we let $P_{S_1,S_2} = P_{\tilde{S}_1,\tilde{S}_2} = P_{S_1}P_{S_2}$. Define two independent random variables $V_1 \in \mathcal{X}_1$ and $V_2 \in \mathcal{X}_2$, whose joint probability distribution $P_{V_1}P_{V_2}$ achieves the rate pair $(I(V_1;Y_2|V_2), I(V_2;Y_1|V_1))$ in Shannon's capacity inner bound. For $j = 1, 2$, we let $\hat{S}_j$ denote the reconstruction variable in the standard RD function of S_j in (1) and choose $P_{\hat{S}_j|S_j}$ that attains $R^{(j)}(D_j)$. Also, we define $(V'_1, V'_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ with $P_{V'_1}P_{V'_2} = P_{V_1}P_{V_2}$ and define $\hat{S}'_j \in \hat{\mathcal{S}}_j$ as the reconstruction variable in the standard RD function of $\tilde{S}_j$ at distortion level D_j, i.e., we set $P_{\hat{S}'_j|\tilde{S}_j} = P_{\hat{S}_j|S_j}$. For $j = 1, 2$, let $U_j \triangleq (V_j, \hat{S}_j)$ and $\tilde{U}_j \triangleq (V'_j, \hat{S}'_j)$ and set $P_{U_j|S_j} = P_{V_j}P_{\hat{S}_j|S_j}$ and $P_{\tilde{U}_j|\tilde{S}_j} = P_{V'_j}P_{\hat{S}'_j|\tilde{S}_j}$. Clearly, the necessary condition in (11) is satisfied. Moreover, set

$$X_j = f_j(\tilde{U}_j, \tilde{S}_j) = f_j((V'_j, \hat{S}'_j), \tilde{S}_j) = V'_j$$

and choose the decoding function g_j as

$$\begin{aligned}\hat{\tilde{S}}_{j'} &= g_j(\tilde{U}_{j'}, \tilde{U}_j, \tilde{S}_j, \tilde{Y}_j)\\ &= g_j((V'_{j'}, \hat{S}'_{j'}), (V'_j, \hat{S}'_j), \tilde{S}_j, \tilde{Y}_j) = \hat{S}'_{j'},\end{aligned}$$

which yields $\mathbb{E}[d_j(\tilde{S}_j, \hat{\tilde{S}}_j)] \le D_j$ for $j = 1, 2$. The above construction ensures that the tuple

$$(\underbrace{P_{V_1}P_{\hat{S}_1|S_1}}_{=P_{U_1|S_1}}, \underbrace{P_{V_2}P_{\hat{S}_2|S_2}}_{=P_{U_2|S_2}}, \underbrace{P_{\tilde{S}_1}P_{\tilde{S}_2}P_{V'_1}P_{\hat{S}'_1|\tilde{S}_1}P_{V'_2}P_{\hat{S}'_2|\tilde{S}_2}}_{=P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}}, P_{\tilde{W}_1,\tilde{W}_2}, f_1, f_2, g_1, g_2)$$

is a configuration in $\Pi'_Z(D_1, D_2)$. Next, using the fact that S_1 and S_2 are independent, one can simplify the sufficient conditions in (15) as follows (the details are given in Appendix B):

$$\begin{aligned}R^{(1)}(D_1) &< I(X_1;Y_2|X_2)\\ R^{(2)}(D_2) &< I(X_1;Y_1|X_2)\end{aligned}$$

which is the achievability result for the SSCC scheme based on the standard lossy source coding and Shannon's random channel coding (without time-sharing).

(iii) *SSCC for the Lossy Transmission of Correlated Sources*: For $j = 1, 2$, we define pairs $(V_1, V_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ and $(V'_1, V'_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ in the same way as in the special case (ii); set the two pairs to have identical distributions, i.e., $P_{V_1}P_{V_2} = P_{V'_1}P_{V'_2}$. Letting $T_j \in \mathcal{T}_j$ denote the auxiliary random variable in the WZ RD function of S_j in (2), we choose $P_{T_j|S_j}$ and the associated decoding function $h_{j'}(T_j, S_{j'})$ that achieves $R^{(j)}_{\mathrm{WZ}}(D_j)$. Similarly, we use $T'_j \in \mathcal{T}_j$ in the WZ RD function of $\tilde{S}_j$ and set $P_{T'_j|\tilde{S}_j} = P_{T_j|S_j}$. Letting $U_j \triangleq (V_j, T_j)$ and $\tilde{U}_j \triangleq (V'_j, T'_j)$, we set $P_{U_j|S_j} = P_{V_j}P_{T_j|S_j}$ and $P_{\tilde{U}_j|\tilde{S}_j} = P_{V'_j}P_{T'_j|\tilde{S}_j}$. Also, set $P_{\tilde{S}_1,\tilde{S}_2} = P_{S_1,S_2}$. Thus, (10) and (11) are satisfied. Moreover, we set the encoding and decoding functions as

$$X_j = f_j(\tilde{U}_j, \tilde{S}_j) = f_j((V'_j, T'_j), \tilde{S}_j) = V'_j$$

and

$$\begin{aligned}\hat{\tilde{S}}_j &= g_{j'}(\tilde{U}_j, \tilde{U}_{j'}, \tilde{S}_{j'}, \tilde{Y}_{j'})\\ &= g_{j'}((V'_j, T'_j), (V'_{j'}, T'_{j'}), \tilde{S}_{j'}, \tilde{Y}_{j'}) = h_{j'}(T'_j, \tilde{S}_{j'}),\end{aligned}$$

such that the decoder satisfies $\mathbb{E}[d_j(\tilde{S}_j, \hat{\tilde{S}}_j)] \le D_j$ for $j = 1, 2$. With the above specifications, we next apply (13) to obtain $P_{\tilde{W}_1,\tilde{W}_2}$, yielding the following configuration in $\Pi'_Z(D_1, D_2)$:

$$(\underbrace{P_{V_1}P_{T_1|S_1}}_{=P_{U_1|S_1}}, \underbrace{P_{V_2}P_{T_2|S_2}}_{=P_{U_2|S_2}}, \underbrace{P_{\tilde{S}_1,\tilde{S}_2}P_{V'_1}P_{T'_1|\tilde{S}_1}P_{V'_2}P_{T'_2|\tilde{S}_2}}_{=P_{\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}}, P_{\tilde{W}_1,\tilde{W}_1}, f_1, f_2, h_1, h_2).$$

Furthermore, using the Markov chain relationship: $T'_1 \multimap \tilde{S}_1 \multimap \tilde{S}_2 \multimap T'_2$ and the memoryless property of the channel, one can easily deduce the following two inequalities from (15):

$$\begin{aligned}R^{(1)}_{\mathrm{WZ}}(D_1) &< I(X_1;Y_2|X_2)\\ R^{(2)}_{\mathrm{WZ}}(D_2) &< I(X_2;Y_1|X_1)\end{aligned}$$

which is the achievability result for the SSCC scheme based on the WZ lossy source coding and Shannon's random channel coding (without time-sharing) [1]. As the derivation is very similar to the previous case (see Appendix B), we omit the details.

(iv) *Correlation-Preserving Coding Scheme for (Almost) Lossless Transmission of Correlated Sources [35]*: Suppose that $\mathcal{S}_j = \hat{\mathcal{S}}_j$ and consider the Hamming distortion measure [22, Sec. 3.6]. We first set $P_{\tilde{S}_1,\tilde{S}_2} = P_{S_1,S_2}$ to meet the necessary condition in (10). Recall the definitions of (V_1, V_2) and (V'_1, V'_2) in the special case (ii) with $P_{V_1}P_{V_2} = P_{V'_1}P_{V'_2}$, which achieve the same rate pair $(I(V_1;Y_2|V_2), I(V_2;Y_1|V_1))$ in Shannon's capacity inner bound. Moreover, we recall the variables $(\hat{S}_1, \hat{S}_2)$ and $(\hat{S}'_1, \hat{S}'_2)$ from the special case (ii), but here we choose $P_{\hat{S}_j|S_j}$ to achieve $R^{(j)}(0)$ in (1) and set $P_{\hat{S}'_j|\tilde{S}_j} = P_{\hat{S}_j|S_j}$ for $j = 1, 2$. Let $U_j \triangleq (V_j, \hat{S}_j)$ and $\tilde{U}_j \triangleq (V'_j, \hat{S}'_j)$, and set $P_{U_j|S_j} = P_{V_j}P_{\hat{S}_j|S_j}$ and $P_{\tilde{U}_j|\tilde{S}_j} = P_{V'_j}P_{\hat{S}'_j|\tilde{S}_j}$. The setting satisfies the condition in (11). We next consider the following encoding and decoding functions:

$$X_j = f_j(\tilde{U}_j, \tilde{S}_j) = f_j((V'_j, \hat{S}'_j), \tilde{S}_j) = V'_j$$

and

$$\begin{aligned}\hat{\tilde{S}}_{j'} &= g_j(\tilde{U}_{j'}, \tilde{U}_j, \tilde{S}_j, Y_j)\\ &= g_j((V'_{j'}, \hat{S}'_{j'}), (V'_j, \hat{S}'_j), \tilde{S}_j, \tilde{Y}_j) = \hat{S}'_{j'}.\end{aligned}$$

Using (13) to obtain $P_{\tilde{W}_1,\tilde{W}_2}$, we ensure that the resulting configuration belongs to $\Pi'_Z(0,0)$. Furthermore, one can easily show that the sufficient conditions in (15) become

$$\begin{aligned} R^{(1)}(0) &= H(\tilde{S}_1|\tilde{S}_2) \\ &< I(V_1';Y_2|V_2',\tilde{S}_2) = I(X_1;Y_2|X_2,\tilde{S}_2) \\ R^{(2)}(0) &= H(\tilde{S}_2|\tilde{S}_1) \\ &< I(V_2';Y_1|V_1',\tilde{S}_1) = I(X_2;Y_1|X_1,\tilde{S}_1) \end{aligned}$$

which recover the achievability conditions in [35, Cor. 8.1] (the rate-one case without coded time-sharing). Note that the block error rate for reconstructing the source messages is asymptotically vanishing here since the above conditions imply that $\lim_{K\to\infty} \Pr(\mathcal{E}) = 0$ (see Appendix A for the definition of the error event $\mathcal{E}$) and hence $\lim_{K\to\infty} \Pr\left((\tilde{S}_j^K, \hat{\tilde{S}}_j^K) \in \mathcal{T}_\epsilon^{(K)}\right) = 1$ for $j = 1,2$, where $\mathcal{T}_\epsilon^{(K)}$ denotes the jointly typical set with parameters K and ϵ as defined in [22]. This result implies that $\lim_{K\to\infty} \Pr\left(\{\tilde{S}_1^K \neq \hat{\tilde{S}}_1^K\} \cup \{\tilde{S}_2^K \neq \hat{\tilde{S}}_2^K\}\right) = 0$.

In fact, since superposition coding is disabled in this simplified scheme, it is unnecessary to use the sliding window decoder. The decoding of each new source block can be done within the same transmission block. The block diagram of such coding system is depicted in Fig. 4 with the following system operations. The source messages S_j^K are first mapped to a digital codeword $U_j^K(M_j)$ with index M_j. The channel inputs X_j^K are then generated via the symbol-by-symbol map $\tilde{f}_j$, which combines the digital information $U_j^K(M_1)$ with the raw (or analog) information S_j^K. Upon receiving Y_j^K, terminal j estimates the codeword index $M_{j'}$ based on all available information. Finally, the decoded codeword $U_{j'}(\hat{M}_{j'})$ and source message S_j^K are passed together through the symbol-by-symbol map $\tilde{g}_j$ to produce $\hat{S}_{j'}^K$. The performance of this specific coding system is analyzed in [2]. The sufficient conditions in the achievability result are identical to those in (15) except that $(\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2)$ are replaced with (S_1,S_2,U_1,U_2).

B. An SSCC Scheme With Adaptive Channel Coding

In the second simplification, we disable superposition coding for the raw source messages; i.e., we let $X_j = F_j(S_j,U_j,\tilde{S}_j,\tilde{U}_j,\tilde{W}_j) \triangleq f_j(U_j,\tilde{U}_j,\tilde{W}_j)$ and $\hat{\tilde{S}}_{j'} = G_j(\tilde{U}_{j'},S_j,U_j,\tilde{S}_j,\tilde{U}_j,\tilde{W}_j,Y_j) \triangleq g_j(\tilde{U}_{j'},\tilde{S}_j)$ for some f_j and g_j, $j=1,2$. Set $P_{\tilde{S}_1,\tilde{S}_2} = P_{S_1,S_2}$ to satisfy (10). Let V_j, $\tilde{V}_j$, and $\tilde{W}_j$ be the auxiliary random variables used in Han's result [15] and let $\gamma_j : \mathcal{V}_j \times \tilde{\mathcal{V}}_j \times \tilde{\mathcal{W}}_j \to \mathcal{X}_j$ denote terminal j's encoding function. Here, we choose $P_{V_1,V_2,\tilde{V}_1,\tilde{V}_2,\tilde{W}_1,\tilde{W}_2}$ and γ_j that achieves the rate pair $(I(\tilde{V}_1;X_2,Y_2,\tilde{V}_2,\tilde{W}_2), I(\tilde{V}_2;X_1,Y_1,\tilde{V}_1,\tilde{W}_1))$ in Han's channel coding inner bound. Note that in Han's result, $P_{V_1,V_2,\tilde{V}_1,\tilde{V}_2,\tilde{W}_1,\tilde{W}_2} = P_{V_1}P_{V_2}P_{\tilde{V}_1}P_{\tilde{V}_2}P_{\tilde{W}_1,\tilde{W}_2|\tilde{V}_1,\tilde{V}_2}$ and $P_{\tilde{V}_j} = P_{V_j}$, $j=1,2$.

Moreover, recall in (2) the auxiliary random variable T_j in the WZ-RD function for S_j, $j=1,2$; we choose $P_{T_j|S_j}$ and the associated decoding function $h_{j'}$ that attains $R_{\mathrm{WZ}}^{(j)}(D_j)$. We also define its counterpart $\tilde{T}_j$ for $\tilde{S}_j$ and set $P_{\tilde{T}_j|\tilde{S}_j} = P_{T_j|S_j}$ for $j=1,2$. Let $U_j \triangleq (V_j,T_j)$ and $\tilde{U}_j \triangleq (\tilde{V}_j,\tilde{T}_j)$ and set $P_{U_j|S_j} = P_{V_j}P_{T_j|S_j}$ and $P_{\tilde{U}_j|\tilde{S}_j} = P_{\tilde{V}_j}P_{\tilde{T}_j|\tilde{S}_j}$, which satisfy (11). Next, we consider the following encoding and decoding functions: $f_j(U_j,\tilde{U}_j,\tilde{W}_j) = \gamma_j(V_j,\tilde{V}_j,\tilde{W}_j)$ and $g_j(\tilde{U}_{j'},\tilde{S}_j) = h_j(\tilde{T}_{j'},\tilde{S}_j)$, which ensures that $\mathbb{E}[d_j(\tilde{S}_j,\hat{\tilde{S}}_j)] \le D_j$ for $j=1,2$. Under the above setting, the joint probability distribution of all involved random variables is then given by

$$\begin{aligned} P_Z = P_{S_1,S_2} &\underbrace{P_{V_1}P_{T_1|S_1}}_{=P_{U_1|S_1}} \underbrace{P_{V_2}P_{T_2|S_2}}_{=P_{U_2|S_2}} P_{\tilde{S}_1,\tilde{S}_2} \\ &\cdot \underbrace{P_{\tilde{V}_1}P_{\tilde{T}_1|\tilde{S}_1}}_{=P_{\tilde{U}_1|\tilde{S}_1}} \underbrace{P_{\tilde{V}_2}P_{\tilde{T}_2|\tilde{S}_2}}_{=P_{\tilde{U}_2|\tilde{S}_2}} \underbrace{P_{\tilde{W}_1,\tilde{W}_2|\tilde{V}_1,\tilde{V}_2}}_{=P_{\tilde{W}_1,\tilde{W}_2|\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2}} \\ &\cdot P_{X_1|V_1,\tilde{V}_1,\tilde{W}_1}P_{X_2|V_2,\tilde{V}_2,\tilde{W}_2}P_{Y_1,Y_2|X_1,X_2}, \end{aligned} \tag{17}$$

where $P_{\tilde{W}_1,\tilde{W}_2|\tilde{V}_1,\tilde{V}_2}$ is specified by Han's result [15] and $P_{X_j|V_j,\tilde{V}_j,\tilde{W}_j}$ is determined by γ_j, $j=1,2$. It can be shown (by definition) that $P_Z = P_Z Q_Z$, thus implying that

$$\begin{aligned} (P_{V_1}P_{T_1|S_1}, P_{V_2}P_{T_2|S_2}, P_{\tilde{S}_1,\tilde{S}_2}P_{\tilde{V}_1}P_{\tilde{T}_1|\tilde{S}_1}P_{\tilde{V}_2}P_{\tilde{T}_2|\tilde{S}_2}, \\ P_{\tilde{W}_1,\tilde{W}_2|\tilde{V}_1,\tilde{V}_2}, \gamma_1,\gamma_2,h_1,h_2) \in \Pi_Z(D_1,D_2). \end{aligned}$$

Letting $\Pi''_Z(D_1,D_2) \subseteq \Pi_Z(D_1,D_2)$ denote the set of all such special configurations, we obtain the following corollary from Theorem 1.

Corollary 2 (SSCC With WZ Source Coding and Han's Adaptive Channel Coding): A distortion pair (D_1,D_2) is achievable for the rate-one lossy transmission of correlated sources over a DM-TWC if there exists a configuration in $\Pi''_Z(D_1,D_2)$ such that

$$R_{\mathrm{WZ}}^{(1)}(D_1) < I(\tilde{V}_1;X_2,Y_2,\tilde{V}_2,\tilde{W}_2), \tag{18a}$$

$$R_{\mathrm{WZ}}^{(2)}(D_2) < I(\tilde{V}_2;X_1,Y_1,\tilde{V}_1,\tilde{W}_1). \tag{18b}$$

Proof: For any configuration in $\Pi''_Z(D_1,D_2)$, the associated stationary distribution P_Z can be factorized into the product form in (17). In addition to the independence between (S_1,S_2,U_1,U_2) and $(\tilde{S}_1,\tilde{S}_2,\tilde{U}_1,\tilde{U}_2,\tilde{W}_1,\tilde{W}_2)$, the quadruple $(\tilde{S}_1,\tilde{S}_2,\tilde{T}_1,\tilde{T}_2)$ is independent of $(\tilde{V}_1,\tilde{V}_2)$. These facts imply the independence between $\tilde{V}_j$ and $(S_{j'},V_{j'},\tilde{S}_{j'},\tilde{T}_{j'})$. Moreover, we have the following Markov chain relationships: $T_1 \multimap\!\!- S_1 \multimap\!\!- S_2 \multimap\!\!- T_2$, $\tilde{T}_1 \multimap\!\!- \tilde{S}_1 \multimap\!\!- \tilde{S}_2 \multimap\!\!- \tilde{T}_2$, and $\tilde{T}_j \multimap\!\!- (\tilde{V}_j,S_{j'},U_{j'},\tilde{S}_{j'},\tilde{T}_{j'}) \multimap\!\!- (\tilde{V}_{j'},\tilde{W}_{j'},X_{j'},Y_{j'})$, $j=1,2$. We now show that (12a) reduces to (18a), where (19), shown at the bottom of the next page, holds since $I(\tilde{U}_1;\tilde{S}_2,\tilde{T}_2|S_2,U_2) = I(\tilde{U}_1;\tilde{S}_2,\tilde{T}_2)$ and $\tilde{U}_j = (\tilde{V}_j,\tilde{T}_j)$, and we have the equivalence in (20), shown at the bottom of the next page, since

$$\begin{aligned} &I(\tilde{S}_1;\tilde{T}_1) - I(\tilde{V}_1,\tilde{T}_1;\tilde{S}_2,\tilde{T}_2) \\ &= I(\tilde{S}_1;\tilde{T}_1) - I(\tilde{T}_1;\tilde{S}_2,\tilde{T}_2) - \underbrace{I(\tilde{V}_1;\tilde{S}_2,\tilde{T}_2|\tilde{T}_1)}_{=0} \\ &= I(\tilde{S}_1;\tilde{T}_1) - I(\tilde{T}_1;\tilde{S}_2,\tilde{T}_2) - I(\tilde{S}_1;\tilde{T}_1|\tilde{S}_2) + I(\tilde{S}_1;\tilde{T}_1|\tilde{S}_2) \\ &= H(\tilde{T}_1) - H(\tilde{T}_1|\tilde{S}_1) - H(\tilde{T}_1) + \underbrace{H(\tilde{T}_1|\tilde{S}_2,\tilde{T}_2)}_{=H(\tilde{T}_1|\tilde{S}_2)} \\ &\qquad - H(\tilde{T}_1|\tilde{S}_2) + \underbrace{H(\tilde{T}_1|\tilde{S}_1,\tilde{S}_2)}_{=H(\tilde{T}_1|\tilde{S}_1)} + I(\tilde{S}_1;\tilde{T}_1|\tilde{S}_2) \\ &= I(\tilde{S}_1;\tilde{T}_1|\tilde{S}_2), \end{aligned}$$

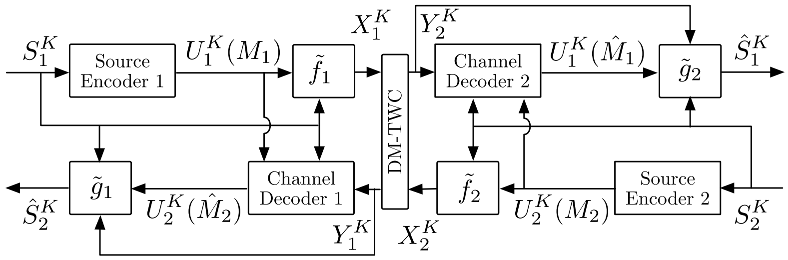


Fig. 4. Rate-one non-adaptive hybrid coding scheme for the transmission of correlated sources over DM-TWCs.

and

$$
\begin{aligned}
&I(\tilde{V}_1,\tilde{T}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2,\tilde{S}_2,\tilde{T}_2)\\
&\quad= I(\tilde{V}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2,\tilde{S}_2,\tilde{T}_2)\\
&\qquad+ \underbrace{I(\tilde{T}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2,\tilde{S}_2,\tilde{T}_2,\tilde{V}_1)}_{=0}\\
&\quad= H(\tilde{V}_1|S_2,U_2,\tilde{S}_2,\tilde{T}_2)\\
&\qquad- H(\tilde{V}_1|S_2,U_2,\tilde{S}_2,\tilde{T}_2,\tilde{V}_2,\tilde{W}_2,X_2,Y_2)\\
&\quad= H(\tilde{V}_1) - H(\tilde{V}_1|\tilde{V}_2,\tilde{W}_2,X_2,Y_2) \qquad (21)\\
&\quad= I(\tilde{V}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2),
\end{aligned}
$$

where (21) holds since $\tilde{V}_1$ is independent of $(S_2,V_2,\tilde{S}_2,\tilde{T}_2)$ given $(\tilde{V}_2,\tilde{W}_2,X_2,Y_2)$. By symmetry, one can also deduce (18b) from (12b), thus completing the proof. ■

We note that by working with super-symbols, we obtain a rate-K/N extension of Corollary 2.

Corollary 3 (General Rate SSCC With WZ Source Coding and Han's Adaptive Channel Coding): A distortion pair (D_1,D_2) is achievable for the rate-K/N lossy transmission of correlated sources over a DM-TWC if

$$K\cdot R_{\mathrm{WZ}}^{(1)}(D_1) < N\cdot I(\tilde{V}_1;X_2,Y_2,\tilde{V}_2,\tilde{W}_2), \qquad (22a)$$

$$K\cdot R_{\mathrm{WZ}}^{(2)}(D_2) < N\cdot I(\tilde{V}_2;X_1,Y_1,\tilde{V}_1,\tilde{W}_1), \qquad (22b)$$

for some joint probability distribution $P_{\tilde{V}_1,\tilde{V}_2,\tilde{W}_1,\tilde{W}_2,X_1,X_2}$ as defined in [15, Section IV].

As Han's channel coding result subsumes Shannon's result, the following corollary is immediate, which is perhaps the simplest SSCC result for our problem setup.

Corollary 4 (General Rate SSCC With WZ Source Coding and Non-Adaptive Channel Coding): A distortion pair (D_1,D_2) is achievable for the rate-K/N lossy transmission of correlated sources over a DM-TWC if

$$K\cdot R_{\mathrm{WZ}}^{(1)}(D_1) < N\cdot I(X_1;Y_2|X_2), \qquad (23a)$$

$$K\cdot R_{\mathrm{WZ}}^{(2)}(D_2) < N\cdot I(X_2;Y_1|X_1), \qquad (23b)$$

for some $P_{X_1}P_{X_2}$.

As our general JSCC scheme (in the proof of Theorem 1) does not consider time-sharing for the sake of simplicity, the channel coding rate pairs obtained by the convex closure operation in Han's and Shannon's inner bound (see Section II-C) are excluded in Corollary 3 and Corollary 4, respectively. However, one can clearly incorporate time-sharing in our coding scheme and Theorem 1. After such convexification operation, any achievable rate pair in Han's (resp., Shannon's) capacity inner bound, i.e., (22) (resp., (23)) will be included. Furthermore, despite the fact that Corollary 3 strictly subsumes Corollary 4, the associated achievable distortion regions are identical when DM-TWCs are symmetric [13]; i.e., when Shannon's inner bound is tight. In such situation, the simpler coding scheme of Corollary 4 is preferred.

To end our discussion on the achievability part, we remark that integrating source refinement and cross-interference mitigation in each channel input is seemingly the most efficient way for source transmission. However, we observe that separately achieving the two objectives can sometimes result in a better performance in related problems such as two-way channel coding [14] and two-way function computation [30]. Thus, it is of interest to ask under what conditions the separation of the two objectives also benefits source-channel communication, which we leave as future research.

V. Converse Results and Complete JSCC Theorems

The last two sections were devoted to the construction of achievable coding schemes. In this section, we derive two outer bounds to the achievable distortion region. Our objective is not only to identify unattainable distortion pairs but also to establish complete JSCC theorems.

A. Two Outer Bounds

Lemmas 1 and 2 provide two outer bounds. Lemma 2 is obtained via a genie-aided argument where the encoder at

$$
\begin{aligned}
& I(\tilde{S}_1;\tilde{U}_1) < I(\tilde{U}_1;S_2,U_2,\tilde{S}_2,\tilde{U}_2,\tilde{W}_2,X_2,Y_2)\\
\Leftrightarrow\quad & I(\tilde{S}_1;\tilde{T}_1) + \underbrace{I(\tilde{S}_1;\tilde{V}_1|\tilde{T}_1)}_{=0} < \underbrace{I(\tilde{U}_1;S_2,U_2)}_{=0} + I(\tilde{U}_1;\tilde{S}_2,\tilde{V}_2,\tilde{T}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2)\\
\Leftrightarrow\quad & I(\tilde{S}_1;\tilde{T}_1) - I(\tilde{U}_1;\tilde{S}_2,\tilde{T}_2|S_2,U_2) < I(\tilde{U}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2,\tilde{S}_2,\tilde{T}_2)\\
\Leftrightarrow\quad & I(\tilde{S}_1;\tilde{T}_1) - I(\tilde{V}_1,\tilde{T}_1;\tilde{S}_2,\tilde{T}_2) < I(\tilde{V}_1,\tilde{T}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2|S_2,U_2,\tilde{S}_2,\tilde{T}_2) \qquad (19)\\
\Leftrightarrow\quad & I(\tilde{S}_1;\tilde{T}_1|\tilde{S}_2) < I(\tilde{V}_1;\tilde{V}_2,\tilde{W}_2,X_2,Y_2) \qquad (20)
\end{aligned}
$$

terminal j can access the decoder side-information $S_{j'}^K$ at terminal j'. The proofs are standard and hence omitted. Details are given in [1] and [2], respectively.

Lemma 1: If a rate-K/N JSCC scheme achieves the distortion levels D_1 and D_2 for the lossy transmission of correlated sources over a DM-TWC, then

$$K \cdot R^{(1)}(D_1) \leq K \cdot I(S_1; S_2) + N \cdot I(X_1; Y_2|X_2), \quad (24a)$$
$$K \cdot R^{(2)}(D_2) \leq K \cdot I(S_1; S_2) + N \cdot I(X_2; Y_1|X_1), \quad (24b)$$

for some P_{X_1,X_2}.

Lemma 2 (Genie-Aided Outer Bound): If a rate-K/N JSCC scheme achieves the distortion levels D_1 and D_2 for the lossy transmission of correlated sources over a DM-TWC, then we have

$$K \cdot R_{S_1|S_2}(D_1) \leq N \cdot I(X_1; Y_2|X_2), \quad (25a)$$
$$K \cdot R_{S_2|S_1}(D_2) \leq N \cdot I(X_2; Y_1|X_1), \quad (25b)$$

for some P_{X_1,X_2}.

Lemmas 1 and 2 generally give different outer bounds; however, the regions are identical for independent sources S_1 and S_2 since in this case $I(S_1; S_2) = 0$ and $R^{(j)}(D_j) = R_{S_j|S_{j'}}(D_j)$. The conditions in (1) and (2) are also equivalent for arbitrarily correlated sources for the specific distortion requirement $(D_1, D_2) = (0, 0)$ since $R_{S_j|S_{j'}}(0) = R^{(j)}(0) - I(S_1; S_2) = H(S_j|S_{j'})$.

B. Complete JSCC Theorems

Matching the achievability results in Section IV with the converse results in Lemmas 1 and 2, we obtain three complete JSCC theorems (Theorems 2-4). We also establish a complete theorem (Theorem 5) for correlated source pairs that have common parts. In the results below, a "symmetric DM-TWC" is a DM-TWC that possesses the symmetry properties defined in [13]. With these properties, Shannon's inner bound in (4) is tight and hence the capacity region is achieved via independent inputs. Moreover, taking the convex closure in (4) is not needed.

Theorem 2 (Lossy Transmission of Independent Sources): For the rate-K/N lossy transmission of independent sources over a symmetric DM-TWC, a distortion pair (D_1, D_2) is achievable if and only if

$$K \cdot R^{(1)}(D_1) \leq N \cdot I(X_1; Y_2|X_2),$$
$$K \cdot R^{(2)}(D_2) \leq N \cdot I(X_2; Y_1|X_1),$$

for some $P_{X_1}P_{X_2}$.

Proof: This result is due to the special case (ii) of Corollary 1 and Lemma 1, together with the facts that $R_{\mathrm{WZ}}^{(j)}(D_j) = R^{(j)}(D_j)$ and $I(S_1; S_2) = 0$ for independent sources pair. ■

Theorem 3 (Almost Lossless Transmission of Correlated Sources): For the rate-K/N transmission of correlated sources over a symmetric DM-TWC, the almost lossless transmission is achievable if and only if

$$K \cdot H(S_1|S_2) \leq N \cdot I(X_1; Y_2|X_2),$$
$$K \cdot H(S_2|S_1) \leq N \cdot I(X_2; Y_1|X_1),$$

for some $P_{X_1}P_{X_2}$.

Proof: In Lemma 1, we have that $K \cdot R^{(j)}(0) - K \cdot I(S_1; S_2) = K \cdot H(S_j|S_{j'})$. Combining this result with the special case (iv) of Corollary 1 then completes the proof. ■

Theorem 4 (Lossy Transmission of Correlated Sources With Equal WZ and Conditional RD Functions): For the rate-K/N lossy transmission of correlated sources whose WZ-RD functions equal to their conditional RD functions over a symmetric DM-TWC, a distortion pair (D_1, D_2) is achievable if and only if

$$K \cdot R_{S_1|S_2}(D_1) \leq N \cdot I(X_1; Y_2|X_2),$$
$$K \cdot R_{S_2|S_1}(D_2) \leq N \cdot I(X_2; Y_1|X_1),$$

for some $P_{X_1}P_{X_2}$.

Proof: The result follows from the special case (iii) of Corollary 1 and Lemma 2. ■

Theorem 5 (Lossy Transmission of Correlated Sources With a Common Part): Assume that correlated sources S_1 and S_2 have a common part S_0 in the sense of Gács-Körner-Witsenhausen and the triplet (S_0, S_1, S_2) forms a Markov chain $S_1 \multimap S_0 \multimap S_2$. For the rate-$K/N$ lossy transmission of such correlated sources over a symmetric DM-TWC, a distortion pair (D_1, D_2) is achievable if and only if

$$K \cdot R_{S_1|S_0}(D_1) \leq N \cdot I(X_1; Y_2|X_2), \quad (26a)$$
$$K \cdot R_{S_2|S_0}(D_2) \leq N \cdot I(X_2; Y_1|X_1), \quad (26b)$$

for some $P_{X_1}P_{X_2}$.

Proof: We construct a two-way coding scheme using two one-way SSCC schemes, one for each direction of the bi-directional transmission. Specifically, we employ the source coding scheme that achieves the distortion level D_j of the conditional RD function $R_{S_j|S_0}^{(j)}(D_j)$ given in (3), $j = 1, 2$, followed by Shannon's one-way channel coding for data protection. The sufficient conditions for achieving the distortion pair (D_1, D_2) as shown in (26) are thus immediate. Note that in this two-way coding scheme, we do not employ time-sharing and the channel inputs X_1 and X_2 are independent.

The proof of the converse part is presented in Appendix C. Although the inputs X_1 and X_2 are arbitrarily correlated in the outer bound result, we can restrict to independent inputs without changing the outer bound region due to the channel symmetry property, i.e., the capacity region of the DM-TWC can be determined via independent channel inputs. Combining this fact with the achievability result then completes the proof. ■

From the above results, one may find that one challenge in obtaining a complete JSCC theorem lies in determining the capacity region for DM-TWCs. Nevertheless, even if the capacity region of a DM-TWC can be exactly determined, obtaining a complete theorem for coding correlated sources is still difficult since the separation principle may not hold (i.e., separate source and channel coding is not optimal) in general. A similar difficulty has been observed in the simpler (one-way) problem of sending correlated sources over multiple access channels [Section 14, 22]. Our JSCC theorems in this section are special instances where the separation principle holds.

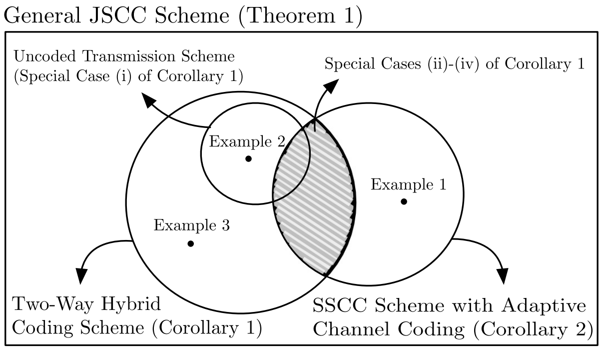


Fig. 5. A general Venn diagram of the achievable distortion regions for the coding schemes presented in Sections III and IV, for a fixed source pair and channel. Moreover, Examples 1-3 in Section VI-A show that certain inclusion relationships can be strict.

VI. Examples and Discussion

In this section, we illustrate our achievability results and discuss possible extensions. The Venn diagram in Fig. 5 summarizes the relationship of the achievable rate regions for the coding schemes in Sections III and IV. We begin with three examples showing that some inclusion relationships can be strict, followed by illustrative examples for Theorems 1, 4, and 5.

A. Examples

Examples 1 and 2 below show that Theorem 1 strictly generalizes Corollary 1 and Corollary 2, respectively. Example 3 not only illustrates a special use of the two-way hybrid coding scheme but also reveals that Corollary 1 strictly subsumes all of its special cases; see Section IV-A. Example 4 shows how a simple instance of our adaptive JSCC helps source transmission. At the end of this section, we provide two examples (Examples 5-6) for Theorem 4 and an example (Example 7) for Theorem 5. Note that except for the Gaussian case examined in Example 6, the Hamming distortion is considered in all examples. Let $\mathrm{Ber}(p)$ denote a Bernoulli random variable with probability of success $p \in [0,1]$, and let $H_{\mathrm{b}}(\cdot)$ denotes the binary entropy function. We will also need the following specialized converse result in Examples 1 and 4, whose proof is similar to Lemma 1.

Proposition 1: Assume that the non-adaptive encoder $f_j : \mathcal{S}_j^K \to \mathcal{X}_j^K$ is used for $j = 1, 2$. If a distortion pair (D_1, D_2) is achievable for the rate-one lossy transmission of independent sources over a DM-TWC, then

$$\begin{aligned} R^{(1)}(D_1) &\le I(X_1; Y_2|X_2, Q),\\ R^{(2)}(D_2) &\le I(X_2; Y_1|X_1, Q), \end{aligned}$$

for some $P_Q P_{X_1|Q} P_{X_2|Q}$.

Note that the pair $(I(X_1; Y_2|X_2, Q), I(X_2; Y_1|X_1, Q))$ under the distribution $P_Q P_{X_1|Q} P_{X_2|Q}$ in Proposition 1 is an alternative expression for the achievable rate pair in Shannon's inner bound (see (4)).

Example 1 (Transmitting Independent Binary Non-Uniform Sources over Dueck's DM-TWC [46]): Consider the independent sources $S_1 = \mathrm{Ber}(0.89)$ and $S_2 = \mathrm{Ber}(0.89)$ so that $H(S_1) = H(S_2) \approx 0.5$. We recall Dueck's DM-TWC [46], where $\boldsymbol{X}_j = (X_{j,1}, X_{j,2})$,[7] $\boldsymbol{Y}_j = (X_{1,1} \cdot X_{2,1}, N_j \oplus X_{j',2}, N_{j'})$, the symbol $\oplus$ denotes the modulo-2 addition, and $N_1 = \mathrm{Ber}(0.5)$ and $N_2 = \mathrm{Ber}(0.5)$ are independent channel noise variables that are independent of all channel inputs and sources. Han [15] showed that the channel coding rate pair $(R_{\mathrm{c},1}, R_{\mathrm{c},2}) = (0.5, 0.5)$ is not achievable via Shannon's random coding scheme but can be achieved via his adaptive channel coding scheme. Based on this fact and Proposition 1, we conclude that the hybrid coding scheme of Corollary 1 cannot achieve the distortion pair $(D_1, D_2) = (0, 0)$ (since it uses non-adaptive encoders and violates the necessary conditions in Proposition 1). By contrast, Corollary 2 shows that the distortion pair $(0, 0)$ is achievable via our general JSCC scheme as $R_{\mathrm{WZ},j}(0) = H(S_j) < R_{\mathrm{c},j}$ holds for $j = 1, 2$. Thus, Theorem 1 strictly subsumes Corollary 1.

Example 2 (Transmitting Correlated Binary Sources over Binary-Multiplying DM-TWCs [4]): Consider the binary-multiplying TWC given by $Y_j = X_1 \cdot X_2$ for $j = 1, 2$. The capacity region of the channel is not known, but it is known that any symmetric achievable channel coding rate pair is component-wise upper bounded by $(0.646, 0.646)$ [9]. Suppose that we want to exchange binary correlated sources with joint probability distribution $P_{S_1,S_2}(0, 0) = 0$ and $P_{S_1,S_2}(s_1, s_2) = 1/3$ for $(s_1, s_2) \ne (0, 0)$. The WZ coding theorem indicates that the minimum source coding rate pair is $(H(S_1|S_2), H(S_2|S_1)) = (0.667, 0.667)$ to achieve the distortion pair $(D_1, D_2) = (0, 0)$. Clearly, this pair is not achievable by *any* SSCC scheme, including the adaptive coding scheme of Corollary 2, because the source coding rate exceeds the largest possible transmission rate for reliable communication. However, the uncoded scheme: $X_j = S_j$ for $j = 1, 2$ can be easily shown to provide lossless transmission. As Corollary 2 and the uncoded scheme are special cases of our general JSCC method, Theorem 1 strictly subsumes Corollary 2.

Example 3 (Transmitting Correlated Binary Sources over a Mixed-Type DM-TWC): Suppose that all alphabets are binary. Let the source messages S_1 and S_2 have the joint probability distribution $P_{S_1,S_2}(1, 0) = 0$ and $P_{S_1,S_2}(s_1, s_2) = 1/3$ for $(s_1, s_2) \ne (1, 0)$. Consider the DM-TWC described by $Y_1 = X_1 \oplus X_2 \oplus N_1$ and $Y_2 = X_1 \cdot X_2$, where $N_1 = \mathrm{Ber}(0.05)$ that is independent of S_j's and X_j's. In other words, we have a (one-way) binary-multiplying channel in one direction and a binary additive channel with additive noise in another direction.

For this channel, none of the special cases of Corollary 1 can achieve the distortion pair $(D_1, D_2) = (0, 0)$. More specifically, the SSCC schemes in the special cases cannot attain the distortion pair since $H(S_1|S_2) < I(X_1; Y_2|X_2)$ and $H(S_2|S_1) < I(X_2; Y_1|X_1)$ cannot hold simultaneously. Moreover, using uncoded transmission in both directions yields the distortion pair $(D_1, D_2) = (0, 0.033)$. However, we can use the two-way hybrid coding scheme in Corollary 1 in the following way: use uncoded transmission from terminal 1 to 2 and use the concatenation of WZ source coding and Shannon's channel coding for the reverse direction. Then

[7] As Dueck's DM-TWC has $\mathcal{X}_j = \{0, 1\}^2$ and $\mathcal{Y}_j = \{0, 1\}^3$, we here use $(X_{j,1}, X_{j,2}) \in \mathcal{X}_j$ to denote the two channel inputs of terminal j.

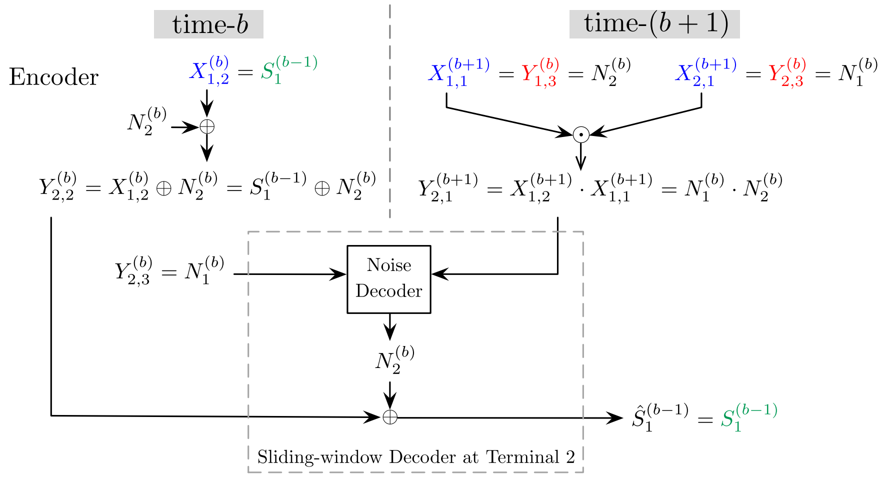


Fig. 6. An illustration of adaptive encoding and sliding-window decoding in Example 4. At time-b, terminal 2 cannot perfectly decode $S_1^{(b-1)}$ from $Y_{2,2}^{(b)}$ due to the additive noise $N_2^{(b)}$. However, at time-$(b+1)$, the adaptive channel inputs $X_{1,1}^{(b+1)}$ and $X_{2,1}^{(b+1)}$ enable a perfect decoding for $N_2^{(b)}$ (based on $Y_{2,3}^{(b)}$, $Y_{2,1}^{(b+1)}$, and the noise correlation) at terminal 2, which can be used to eliminate the noise in $Y_{2,2}^{(b)}$ and achieve error-free transmission.

the distortion pair $(0,0)$ is achievable. This example shows that Corollary 1 is a strictly generalization of its presented special cases.

Example 4 (Transmitting Independent Binary Uniform Sources over Dueck's DM-TWC): Consider the almost lossless transmission of the independent sources $S_1 = \mathrm{Ber}(0.5)$ and $S_2 = \mathrm{Ber}(0.5)$ through Dueck's DM-TWC (given in Example 1). Here, the binary noise variables N_1 and N_2 are assumed to be correlated with joint distribution given by $P_{N_1,N_2}(0,0) = 0$ and $P_{N_1,N_2}(n_1,n_2) = 1/3$ for $(n_1,n_2) \neq (0,0)$. For this channel, the optimal symmetric rate pair in Proposition 1 is obtained as $(I(X_1;Y_2|X_2), I(X_2;Y_1|X_1)) = (0.9503, 0.9503)$. Since the required source coding rate $R_{\mathrm{WZ}}^{(j)}(0) = H(S_j) = 1$ (at terminal j) exceeds the outer bound in Proposition 1, the hybrid coding scheme in Corollary 1 cannot achieve the distortion pair $(D_1,D_2) = (0,0)$.

By contrast, the following use of our general JSCC scheme provides rate-one lossless transmission. Suppose that we exchange a length-K of such source pair via $K+1$ channel uses. Clearly, the transmission rate approaches one as K goes to infinity. For $j = 1,2$, we next set $(X_{j,1}^{(1)}, X_{j,2}^{(1)}) = (1, S_j^{(1)})$, $(X_{j,1}^{(K+1)}, X_{j,2}^{(K+1)}) = (Y_{j,3}^{(K)}, 1)$, and $(X_{j,1}^{(b)}, X_{j,2}^{(b)}) = (Y_{j,3}^{(n-1)}, S_j^{(b)})$ for $b = 2,3,\ldots,K$, where the superscripts represent time index. Via such adaptive encoding, terminal j can exploit the correlation between N_1 and N_2 to perfectly decode $N_j^{(b-1)}$ from $Y_{j,1}^{(b)}$ and $Y_{j,3}^{(b-1)}$ and reconstruct $S_{j'}^{(b-1)}$ as $\hat{S}_{j'}^{(b-1)} = N_j^{(b-1)} \oplus Y_{j,2}^{(b-1)} = S_{j'}^{(b-1)}$ for all $2 \leq b \leq K+1$, thus achieving zero-error transmission. For $2 \leq b \leq K$, the above encoding and decoding procedure is depicted in Fig. 6. Note that whether or not the SSCC scheme in Corollary 2 achieves the same performance remains unclear.

Example 5 (Transmitting Binary Correlated Sources With Z-channel Correlation over Binary Additive Noise DM-TWCs): Suppose that all alphabets are binary. Given $0 \leq \epsilon_1, \epsilon_2 < 0.5$, the binary additive noise DM-TWC is described by $Y_j = X_j \oplus X_{j'} \oplus N_j$, $j = 1,2$, where the channel noise variables $N_1 = \mathrm{Ber}(\epsilon_1)$ and $N_2 = \mathrm{Ber}(\epsilon_2)$ are independent of each other, of the source messages, and of the channel inputs. The capacity region of the channel is given by [45]: $\{(R_{c_1}, R_{c_2}) : 0 \leq R_{c_1} \leq 1 - H_{\mathrm{b}}(\epsilon_2), 0 \leq R_{c_2} \leq 1 - H_{\mathrm{b}}(\epsilon_1)\}$. Consider the binary correlated source pair (S_1,S_2) with Z-channel correlation [44]; i.e., the transition matrices $[P_{S_2|S_1}(\cdot|\cdot)]$ and $[P_{S_1|S_2}(\cdot|\cdot)]$ between the sources S_1 and S_2 can be interpreted as a Z-channel and a reverse Z-channel, respectively. Assume that the crossover probabilities of the Z-type channels are α_1 and α_2, respectively. Let $P_{S_1}(1) = q_1$ and $P_{S_2}(1) = q_2$, where q_2 is a function of q_1 and α_1 (note that one may also write q_1 as a function of q_2 and α_2). According to Theorem 4, the achievable distortion region for the rate-K/N transmission consists of all pairs (D_1,D_2) that satisfy the inequalities, shown at the bottom of the next page.

Example 6 (Transmitting Correlated Gaussian Sources over DM-TWCs With Additive White Gaussian Noise (AWGN) DM-TWCs): Consider the squared-error distortion measure. The AWGN DM-TWC is described by $Y_j = X_j + X_{j'} + N_j$, $j = 1,2$, where N_1 and N_2 are independent zero mean Gaussian noises with variance σ_1^2 and σ_2^2, respectively, and are independent of the source messages and of the channel inputs. The average power of channel inputs X_j is set as P_j for $j = 1,2$. Moreover, the correlated sources S_1 and S_2 are considered to be zero-mean unit-variance jointly Gaussian random variables with correlation coefficient ρ for some $0 \leq \rho \leq 1$. For this setting, Theorem 4 yields the achievable distortion region $\{(D_1,D_2) : D_j \geq (1-\rho^2)(1+\frac{P_j}{\sigma_{j'}})^{\frac{K}{N}}, j = 1,2\}$, for the rate-$K/N$ transmission. The detailed derivation can be found in [1, Lemma 4].

Example 7 (Transmitting Quaternary Correlated Sources Over Binary Additive Noise DM-TWCs): Suppose that $\mathcal{S}_1 = \mathcal{S}_2 = \hat{\mathcal{S}}_1 = \hat{\mathcal{S}}_2 = \{A,B,C,D\}$ and $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_1 = \mathcal{Y}_2 = \{0,1\}$. Consider the correlated source pair with joint probability distribution given by

$$P_{S_1,S_2}(s_1,s_2) = \begin{cases} \frac{1}{8} & \text{if } (s_1,s_2) \in \{A,B\} \times \{A,B\} \\ & \qquad \cup \{C,D\} \times \{C,D\}, \\ 0 & \text{otherwise.} \end{cases}$$

For such sources, we observe a binary common part S_0; $S_0 = 0$ and $S_0 = 1$ are corresponding to $S_1, S_2 \in \{A, B\}$ and $S_1, S_2 \in \{C, D\}$, respectively. Given this common part, we can decompose S_j into (S_0, S'_j), where $S'_j = \mathrm{Ber}(0.5)$. It is easy to show that S_j and (S_0, S'_j) have a one-to-one correspondence and the Markov chain relationship $S'_1 \multimap S_0 \multimap S'_2$ holds. Moreover, the conditional RD function $P_{S'_j|S_0}(D_j)$ is given by $P_{S'_j|S_0}(D_j) = 1 - H_\mathrm{b}(D_j)$ for $0 \le D_j \le 0.5$.

Due to the above decomposition, the terminals only need to exchange (S'_1, S'_2). When transmitting the pair (S'_1, S'_2) over the binary additive noise DM-TWCs (defined in Example 5) at rate-K/N, we can apply Theorem 5 to characterize the achievable distortion region of the overall system, which is the convex hull of all distortion pairs (D_1, D_2) satisfying

$$K(1 - H_\mathrm{b}(D_1)) \le N(1 - H_\mathrm{b}(\epsilon_2)),$$
$$K(1 - H_\mathrm{b}(D_2)) \le N(1 - H_\mathrm{b}(\epsilon_1)).$$

B. Adaptive Coding With More Past Information

In our JSCC scheme (detailed in Appendix A), we merely use the most recent channel inputs and outputs $(X_j^{(t-1)}, Y_j^{(t-1)})$ to generate the current channel input $X_j^{(t)}$. Although ideally one would use the entire past channel input and output history for adaptive coding, the accumulated information in this case causes the Markov chain not only to have a time-varying transition kernel but also to drastically expand the state space. The idea to jointly optimize the terminals' transmission via a stationary Markov chain becomes infeasible. In the following, we sketch two coding strategies to deal with this problem. Each of the strategies can be directly integrated into our JSCC scheme, but the encoding/decoding complexity will be higher and the sufficient conditions will be significantly more complicated than the current ones.

The first strategy is to generate $X_j^{(t)}$ as a function of the past μ channel inputs $(X_j^{(t-\mu)}, X_j^{(t-\mu+1)}, \ldots, X_j^{(t-1)})$ and outputs $(Y_j^{(t-\mu)}, Y_j^{(t-\mu+1)}, \ldots, Y_j^{(t-1)})$ for some $\mu > 1$, which is similar to the memory-μ channel coding for DM-TWCs [18, Section 4.4]. This strategy increases the encoding and decoding complexity, but the state space complexity of the Markov chain is constant.

The second strategy quantizes the past channel inputs and outputs at each terminal into a set with fixed size. The channel inputs can be then generated as a function of the quantized information in that set, rather than the entire past information. This strategy is similar to the Q-graph channel coding for single-output DM-TWCs [19], and it adds a minor encoding cost. However, as the quantized knowledge is not necessarily a sufficient statistic for optimal decoding, we still need to store all past information, which clearly increases system complexity.[8]

C. Adaptive Coding With Incremental Side-Information

Our adaptive coding mainly coordinates the terminals' transmission on the shared channel as we did not attempt to apply Kaspi's interactive source coding idea [21] to make the best use of the sequentially received signals. Here, we give an SSCC scheme that encompasses both ideas.

The exchange of correlated sources S_1^K and S_2^K is now accomplished in L rounds for some $L \ge 1$, which comprises N channel uses (note that N is a function of K). Specifically, for $1 \le l \le L$, let N_l denote the number of channel uses in the lth round of transmission, where $\sum_{l=1}^L N_l = N$. In each round, viewing the previously transmitted and decoded source codewords as side-information, each terminal applies binning for source coding, followed by Han's adaptive channel coding. Each terminal also decodes the other terminal's source codeword at the end of each transmission round. After L rounds, each terminal reconstructs the other terminal's source messages from the side-information and its own source messages. Clearly, this simple SSCC scheme allows two-way simultaneous transmission and interactive source coding. We summarize the achievability result in Proposition 2 below (without proof). Here, $T_{j,l}$, $j = 1, 2$ and $l = 1, 2, \ldots, L$, are auxiliary random variables.

Proposition 2: A distortion pair (D_1, D_2) is achievable for the rate-K/N lossy transmission of correlated sources over a DM-TWC if for all $1 \le l \le L$, we have that

$$K \cdot I(S_1; T_{1,l}|S_2, T_1^{l-1}, T_2^{l-1}) < N_l \cdot I(\tilde{V}_{1,l}; X_{2,l}, Y_{2,l}, \tilde{V}_{2,l}, \tilde{W}_{2,l}),$$
$$K \cdot I(S_2; T_{2,l}|S_1, T_1^{l-1}, T_2^{l-1}) < N_l \cdot I(\tilde{V}_{2,l}; X_{1,l}, Y_{1,l}, \tilde{V}_{1,l}, \tilde{W}_{1,l}),$$

for some $P_{\tilde{V}_{1,l}, \tilde{V}_{2,l}, \tilde{W}_{1,l}, \tilde{W}_{2,l}, X_{1,l}, X_{2,l}}$ as defined in [15, Section IV] and

$$P_{T_1^L, T_2^L|S_1,S_2} = \prod_{l=1}^{L} P_{T_{1,l}|S_1, T_1^{l-1}, T_2^{l-1}} P_{T_{2,l}|S_2, T_1^{l-1}, T_2^{l-1}}$$

and two decoding functions $\hat{S}_{j'} = g_j(S_j, T_j^L, T_{j'}^L)$ such that $\mathbb{E}[d_j(S_j, \hat{S}_j)] \le D_j$ for $j = 1, 2$.

Note that the above proposition reduces to Corollary 3 when $L = 1$. In light of this, it is of interest to ask if there exists a general adaptive JSCC scheme that integrates both features and subsumes all of our presented achievability results. We leave this question for future research.

[8]One can apply sliding-window decoding to limit the amount of past information at each receiver.

$$K(1 - q_1 + q_1\alpha_1)\left[H_\mathrm{b}\left(\frac{q_1\alpha_1}{1 - q_1 + q_1\alpha_1}\right) - H_\mathrm{b}\left(\frac{D_1}{1 - q_1 + q_1\alpha_1}\right)\right] \le N(1 - H_\mathrm{b}(\epsilon_2)),$$
$$K(1 - q_2 + q_2\alpha_2)\left[H_\mathrm{b}\left(\frac{q_2\alpha_2}{1 - q_2 + q_2\alpha_2}\right) - H_\mathrm{b}\left(\frac{D_2}{1 - q_2 + q_2\alpha_2}\right)\right] \le N(1 - H_\mathrm{b}(\epsilon_1)).$$

VII. Conclusion

We constructed an adaptive coding scheme to prove a direct JSCC theorem, which characterizes an achievable distortion region for two-way lossy simultaneous transmission. Our adaptive coding method demonstrates a way to coordinate the independent transmissions of the terminals; it also underscores the importance of preserving source correlation as illustrated via several examples. Moreover, our coding scheme subsumes several simple non-adaptive coding methods, providing a unified transmission framework that allows for diverse various system complexity and performance trade-offs. Although the general form of our scheme is complex, in many cases its SSCC instances suffice to achieve the optimal performance. Future directions include adaptive coding based on the SSCC structure, symbol-wise adaptive coding (as opposed to block-wise adaptive coding), and practical joint source-channel code design for our problem setup. It is also of interest to refine the outer bounds and derive a complete characterization of the achievable RD region for two-way source-channel communication (in either single-letter or multi-letter expression).

Appendix

A. Proof of Theorem 1

Before presenting the proof, we remark that after specifying our coding method and the set of stationary configurations, one can prove the same result using [43, Theorem 1] since our adaptive coding system involves block-wise operations. Although this proof approach can eliminate certain error analyses, it requires extra ingredients such as "unfolding the network" and determining all coding parameters. This proof approach also provides less insight into our adaptive coding scheme. To better convey the ideas behind our coding scheme and avoid introducing new terminology, we follow a standard argument here.

Let $\mathcal{T}_\epsilon^{(n)}$ denote the typical set of sequences with parameters $n \in \mathbb{Z}_+$ and $\epsilon > 0$ as defined in [22]; the domain of $\mathcal{T}_\epsilon^{(n)}$ will clear from the context and hence omitted. Here, we set $n = N = K$ as we consider the rate-one transmission. For $j = 1, 2$ and $b = 1, 2, \cdots, B$, we define $2^{nR_j^{(b)}}$ as the size of terminal j's codebook $\mathcal{C}_j^{(b)}$, which is used to encode the b-th block $\boldsymbol{S}_j^{(b)}$ of source messages. For an event $\mathcal{E}$, we let $\overline{\mathcal{E}}$ denote its complement.

<u>*Codebook Generation*</u>: Given a configuration in $\Pi_Z(D_1, D_2)$, generate two length-n sequences $(\tilde{\boldsymbol{S}}_1^{(1)}, \tilde{\boldsymbol{S}}_2^{(1)}, \tilde{\boldsymbol{U}}_1^{(1)}, \tilde{\boldsymbol{U}}_2^{(1)}, \tilde{\boldsymbol{W}}_1^{(1)}, \tilde{\boldsymbol{W}}_2^{(1)})$ and $(\boldsymbol{S}_1^{(B+1)}, \boldsymbol{S}_2^{(B+1)}, \boldsymbol{U}_1^{(B+1)}, \boldsymbol{U}_2^{(B+1)})$ to initialize and terminate the $(B+1)$-blocks encoding process with distributions

$$P_{\tilde{\boldsymbol{S}}_1^{(1)}, \tilde{\boldsymbol{S}}_2^{(1)}, \tilde{\boldsymbol{U}}_1^{(1)}, \tilde{\boldsymbol{U}}_2^{(1)}, \tilde{\boldsymbol{W}}_1^{(1)}, \tilde{\boldsymbol{W}}_2^{(1)}}(\tilde{\boldsymbol{s}}_1^{(1)}, \tilde{\boldsymbol{s}}_2^{(1)}, \tilde{\boldsymbol{u}}_1^{(1)}, \tilde{\boldsymbol{u}}_2^{(1)}, \tilde{\boldsymbol{w}}_1^{(1)}, \tilde{\boldsymbol{w}}_2^{(1)}) = \prod_{i=1}^{n} P_{\tilde{S}_1, \tilde{S}_2, \tilde{U}_1, \tilde{U}_2, \tilde{W}_1, \tilde{W}_2}(\tilde{s}_{1,i}^{(1)}, \tilde{s}_{2,i}^{(1)}, \tilde{u}_{1,i}^{(1)}, \tilde{u}_{2,i}^{(1)}, \tilde{w}_{1,i}^{(1)}, \tilde{w}_{2,i}^{(1)}) \tag{27}$$

and

$$P_{\boldsymbol{S}_1^{(B+1)}, \boldsymbol{S}_2^{(B+1)}, \boldsymbol{U}_1^{(B+1)}, \boldsymbol{U}_2^{(B+1)}}(\boldsymbol{s}_1^{(B+1)}, \boldsymbol{s}_2^{(B+1)}, \boldsymbol{u}_1^{(B+1)}, \boldsymbol{u}_2^{(B+1)}) = \prod_{i=1}^{n} P_{S_1, S_2, U_1, U_2}(s_{1,i}^{(B+1)}, s_{2,i}^{(B+1)}, u_{1,i}^{(B+1)}, u_{2,i}^{(B+1)}). \tag{28}$$

Moreover, generate codebooks $\mathcal{C}_j^{(b)} \triangleq \{\boldsymbol{U}_j^{(b)}(m_j^{(b)}) : m_j^{(b)} = 1, 2, \ldots, 2^{nR_j^{(b)}}\}$ for $b = 1, 2, \ldots, B$ and $j = 1, 2$, where $\boldsymbol{U}_j^{(b)}(m_j^{(b)})$ is a length-n sequence distributed according to $P_{\boldsymbol{U}_j}(\boldsymbol{u}_j^{(b)}(m_j^{(b)})) = \prod_{i=1}^{n} P_{U_j}(u_{j,i}^{(b)}(m_j^{(b)}))$ and $\boldsymbol{U}_j^{(b)}(m_j^{(b)})$'s are independent of each other. The initialization and termination sequences and all codebooks are revealed to both terminals. We note that due to the construction of the Markov chain $\{Z^{(t)}\}$, the codebook $\mathcal{C}_j^{(b)}$ is also used for $\tilde{\boldsymbol{U}}_j^{(b+1)}$.

<u>*Encoding*</u>: Let $\epsilon_1 > \epsilon > 0$. For $b = 1, 2, \ldots, B$ and $j = 1, 2$, terminal j finds $m_j^{(b)}$ such that $(\boldsymbol{S}_j^{(b)}, \boldsymbol{U}(m_j^{(b)})) \in \mathcal{T}_{\epsilon_1}^{(n)}$. If there is more than one such index, the encoder chooses one of them at random. If there is no such index, it chooses an index at random from $\{1, 2, \ldots, 2^{nR_j^{(b)}}\}$. The transmitter then sends $\boldsymbol{X}_j^{(b)}$, where $X_{j,i}^{(b)} = F_j(S_{j,i}^{(b)}, U_{j,i}^{(b)}(m_j^{(b)}), \tilde{S}_{j,i}^{(b)}, \tilde{U}_{j,i}^{(b)}, \tilde{W}_{j,i}^{(b)})$ for $i = 1, 2, \ldots, n$, $\tilde{S}_{j,i}^{(b)} = S_{j,i}^{(b-1)}$, $\tilde{U}_{j,i}^{(b)} = U_{j,i}^{(b-1)}$, and $\tilde{W}_{j,i}^{(b)} = (X_{j,i}^{(b-1)}, Y_{j,i}^{(b-1)})$ for $b = 2, 3, \ldots, B$. For $b = B+1$, $\boldsymbol{X}^{(B+1)}$ is generated in the same way using the termination sequence.

<u>*Decoding*</u>: For $b=2, 3, \ldots, B+1$ and $j, j'=1, 2$ with $j \neq j'$, terminal j finds an index $\hat{m}_{j'}^{(b-1)}$ such that $(\boldsymbol{S}_j^{(b)}, \boldsymbol{U}_j^{(b)}, \tilde{\boldsymbol{S}}_j^{(b)}, \tilde{\boldsymbol{U}}_j^{(b)}, \tilde{\boldsymbol{U}}_{j'}^{(b)}(\hat{m}_{j'}^{(b-1)}), \tilde{\boldsymbol{W}}_j^{(b)}, \boldsymbol{X}_j^{(b)}, \boldsymbol{Y}_j^{(b)}) \in \mathcal{T}_\epsilon^{(n)}$, where $\tilde{\boldsymbol{U}}_{j'}^{(b)}(\hat{m}_{j'}^{(b-1)}) \in \mathcal{C}_{j'}^{(b-1)}$. If there is more than one choice, the decoder chooses one of them at random. If there is no such index, it chooses one at random from $\{1, 2, \ldots, 2^{nR_{j'}^{(b)}}\}$. The reconstruction for the source message $\boldsymbol{S}_{j'}^{(b-1)}$ is given by $\hat{S}_{j',i}^{(b-1)} = G_j(\tilde{U}_{j',i}^{(b)}(\hat{m}_{j'}^{(b-1)}), S_{j,i}^{(b)}, U_{j,i}^{(b)}, \tilde{S}_{j,i}^{(b)}, \tilde{U}_{j,i}^{(b)}, \tilde{W}_{j,i}^{(b)}, Y_{j,i}^{(b)})$ for $i = 1, 2, \ldots, n$.

<u>*Performance Analysis*</u>: Let $M_j^{(b)}$ and $\hat{M}_j^{(b)}$ denote the random encoded and decoded indices for $\boldsymbol{S}_j^{(b)}$. We first define the events $\mathcal{E}_1^{(b)}$, $b = 1, 2, \ldots, B+1$, in (29a), shown at the bottom of the next page, for terminal 1. We analogously define the events $\mathcal{E}_2^{(b)}$ for terminal 2 (not shown here) and consider the error event $\mathcal{E} = \cup_{b=1}^{B+1} \mathcal{E}_1^{(b)} \cup \mathcal{E}_2^{(b)}$. The expected distortion of terminal j's source reconstruction (averaged with respect to all codebooks, source messages, channel inputs, and channel outputs) can be bounded by

$$\frac{1}{B} \sum_{b=1}^{B} \mathbb{E}[d_j(\boldsymbol{S}_j^{(b)}, \hat{\boldsymbol{S}}_j^{(b)})]$$
$$\leq \Pr(\mathcal{E}) d_{j,\max} + \frac{1}{B} \sum_{b=1}^{B} \Pr\left(\overline{\mathcal{E}}\right) \mathbb{E}[d_j(\boldsymbol{S}_j^{(b)}, \hat{\boldsymbol{S}}_j^{(b)}) | \overline{\mathcal{E}}] \tag{30}$$
$$\leq \Pr(\mathcal{E}) d_{j,\max} + \frac{1}{B} \sum_{b=1}^{B} (1+\epsilon) \mathbb{E}[d_j(S_j^{(b)}, \hat{S}_j^{(b)})] \tag{31}$$
$$= \Pr(\mathcal{E}) d_{j,\max} + (1+\epsilon) \mathbb{E}[d_j(S_j, \hat{S}_j)] \tag{32}$$
$$\leq \Pr(\mathcal{E}) d_{j,\max} + (1+\epsilon) D_j, \tag{33}$$

where (30) follows from $\mathbb{E}[d_j(\boldsymbol{S}_j^{(b)}, \hat{\boldsymbol{S}}_j^{(b)})|\mathcal{E}] \leq d_{j,\max}$ with $d_{j,\max} \triangleq \max_{s_j,\hat{s}_j} d_j(s_j,\hat{s}_j)$, (31) is due to the typical average lemma [22], (32) follows from the stationarity of the Markov chain, and the last inequality holds by assumption.

If we can further show that $\Pr(\mathcal{E}) \to 0$ and the joint source-channel coding rate goes to one as both n and B go to infinity, then the distortion pair $((1+\epsilon)D_1, (1+\epsilon)D_2)$ is achievable. Note that it suffices to show that $\Pr(\mathcal{E}_j^{(1)}) \to 0$ and $\Pr(\mathcal{E}_j^{(b)} \cap \overline{\mathcal{E}}_j^{(b-1)}) \to 0$ for all $j = 1,2$ and $b = 2,3,\ldots,B+1$ since by the identity $\cup_{b=1}^{B}\mathcal{E}_j^{(b)} = \mathcal{E}_j^{(1)} \cup (\cup_{b=2}^{B} \mathcal{E}_j^{(b)} \cap \overline{\mathcal{E}}_j^{(b-1)})$, we have

$$\Pr(\mathcal{E}) \leq \Pr(\mathcal{E}_1^{(1)}) + \Pr(\mathcal{E}_2^{(1)}) + \sum_{b=2}^{B+1}\Big(\Pr(\mathcal{E}_1^{(b)} \cap \overline{\mathcal{E}}_1^{(b-1)}) + \Pr(\mathcal{E}_2^{(b)} \cap \overline{\mathcal{E}}_2^{(b-1)})\Big).$$

Due to symmetry, we only analyze $\Pr(\mathcal{E}_1^{(1)})$ and $\Pr(\mathcal{E}_1^{(b)} \cap \overline{\mathcal{E}}_1^{(b-1)})$. For $j = 1,2$ and $b = 1,2,\ldots,B+1$, we first define

$$\begin{aligned}
\mathcal{F}_j^{(b)} &= \{(\boldsymbol{S}_j^{(b)}, \boldsymbol{U}_j^{(b)}(m_j^{(b)})) \notin \mathcal{T}_{\epsilon_1}^{(n)} \text{ for all } m_j^{(b)}\},\\
\mathcal{F}_3^{(b)} &= \{(\boldsymbol{S}_1^{(b)}, \boldsymbol{S}_2^{(b)}, \boldsymbol{U}_1^{(b)}(M_1^{(b)}), \boldsymbol{U}_2^{(b)}(M_2^{(b)}),\\
&\qquad \tilde{\boldsymbol{S}}_1^{(b)}, \tilde{\boldsymbol{S}}_2^{(b)}, \tilde{\boldsymbol{U}}_1^{(b)}(M_1^{(b-1)}), \tilde{\boldsymbol{U}}_2^{(b)}(M_2^{(b-1)}),\\
&\qquad \tilde{\boldsymbol{W}}_1^{(b)}, \tilde{\boldsymbol{W}}_2^{(b)} \boldsymbol{X}_1^{(b)}, \boldsymbol{X}_2^{(b)}, \boldsymbol{Y}_1^{(b)}, \boldsymbol{Y}_2^{(b)}) \notin \mathcal{T}_\epsilon^{(n)}\},\\
\mathcal{F}_4^{(b)} &= \{\exists\, \hat{m}_1^{(b-1)} \neq M_1^{(b-1)} \text{ s.t. } (\boldsymbol{S}_2^{(b)}, \boldsymbol{U}_2^{(b)}(M_2^{(b)}), \tilde{\boldsymbol{S}}_2^{(b)},\\
&\qquad \tilde{\boldsymbol{U}}_1^{(b)}(\hat{m}_1^{(b-1)}), \tilde{\boldsymbol{U}}_2^{(b)}(M_2^{(b-1)}), \tilde{\boldsymbol{W}}_2^{(b)}, \boldsymbol{X}_2^{(b)}, \boldsymbol{Y}_2^{(b)}) \in \mathcal{T}_\epsilon^{(n)}\},
\end{aligned}$$

with the exception that $\mathcal{F}_3^{(1)} \triangleq \mathcal{E}_1^{(1)}$ and $\mathcal{F}_3^{(B+1)} \triangleq \mathcal{E}_1^{(B+1)}$ due to the initialization and termination phases of the encoding process. Next, we use the following results (whose proofs are omitted) to obtain (12a).

Claim 1: For $b = 2,3,\ldots,B+1$, the event $\overline{\mathcal{F}}_3^{(b)} \cap \overline{\mathcal{F}}_4^{(b)}$ implies that $\hat{M}_1^{(b-1)} = M_1^{(b-1)}$.

Claim 2: $\mathcal{E}_1^{(1)} \subseteq \mathcal{F}_1^{(1)} \cup \mathcal{F}_2^{(1)} \cup (\overline{\mathcal{F}}_1^{(1)} \cap \overline{\mathcal{F}}_2^{(1)} \cap \mathcal{E}_1^{(1)})$

Claim 3: The inclusion $\mathcal{E}_1^{(b)} \cap \overline{\mathcal{E}}_1^{(b-1)} \subseteq \mathcal{F}_1^{(b)} \cup \mathcal{F}_2^{(b)} \cup (\overline{\mathcal{F}}_1^{(1)} \cap \overline{\mathcal{F}}_2^{(1)} \cap \mathcal{F}_3^{(b)} \cap \overline{\mathcal{E}}_1^{(b-1)}) \cup \mathcal{F}_4^{(b)}$ holds for $b = 2,3,\ldots,B$.

Claim 4: $\mathcal{E}_1^{(B+1)} \cap \overline{\mathcal{E}}_1^{(B)} \subseteq (\mathcal{F}_3^{(B+1)} \cap \overline{\mathcal{E}}_1^{(B)}) \cup \mathcal{F}_4^{(B+1)}$

Claim 5: If $R_j^{(1)} > I(S_j;U_j) + \delta_1(\epsilon_1)$, then we have that $\lim_{n\to\infty} \Pr(\mathcal{E}_j^{(1)}) = 0$.

Claim 6: If $R_1^{(B)} < I(\tilde{U}_1; S_2, U_2, \tilde{S}_2, \tilde{U}_2, \tilde{W}_2, X_2, Y_2) - \delta(\epsilon)$, then $\lim_{n\to\infty} \Pr(\mathcal{E}_1^{(B+1)} \cap \overline{\mathcal{E}}_1^{(B)}) = 0$.

Claim 7: For $b = 2,3,\ldots,B$, if $R_j^{(b)} > I(S_j;U_j) + \delta_1(\epsilon_1)$ and $R_1^{(b-1)} < I(\tilde{U}_1; S_2, U_2, \tilde{S}_2, \tilde{U}_2, \tilde{W}_2, X_2, Y_2) - \delta(\epsilon)$, then $\lim_{n\to\infty} \Pr(\mathcal{E}_1^{(b)} \cap \overline{\mathcal{E}}_1^{(b-1)}) = 0$.

Roughly speaking, the non-negative quantities $\delta_1(\epsilon_1)$ and $\delta(\epsilon)$ above arise from the standard typicality arguments and $\lim_{\epsilon_1\to 0}\delta_1(\epsilon_1) = 0$ and $\lim_{\epsilon\to 0}\delta(\epsilon) = 0$. Here, Claim 1 holds due to the definitions of $\bar{\mathcal{F}}_3^{(b)}$ and $\bar{\mathcal{F}}_4^{(b)}$; Claim 2 holds since the right-hand-side is equal to $\mathcal{E}_1^{(1)} \cup \mathcal{F}_1^{(1)} \cup \mathcal{F}_2^{(1)}$. Claims 3 and 4 are shown using the fact that $\mathcal{E}_1^{(b)} \subseteq \mathcal{F}_3^{(b)} \cup \mathcal{F}_4^{(b)}$, which is a consequence of Claim 1. Claims 5-7 are derived based on Claims 2-4, respectively. More specifically, the union bound is applied to each inclusion relationship (in Claims 2-4) to upper bound the probability of the event on the left-hand-side. A thorough analysis then yields the conditions in Claims 5-7, which ensure that all terms in the upper bound asymptotically vanish. The proofs of Claims 5-7 invoke the covering lemma [22], the conditional typical lemma [22], and [38, Lemma 1].

Swapping the role of terminals 1 and 2, we obtain $\lim_{n\to\infty} \Pr(\mathcal{E}_2^{(1)}) = 0$ and are such that $\lim_{n\to\infty} \Pr(\mathcal{E}_2^{(b)} \cap \overline{\mathcal{E}}_2^{(b-1)}) = 0$ for $b = 2,3,\ldots,B+1$ provided that $R_j^{(b)} > I(S_j;U_j) + \delta_1(\epsilon_1)$ for $j = 1,2$ and $b = 1,2,\ldots,B$ and $R_2^{(b-1)} < I(\tilde{U}_2; S_1, U_1, \tilde{S}_1, \tilde{U}_1, \tilde{W}_1, X_1, Y_1) - \delta(\epsilon)$ for $b = 2,3,\ldots,B+1$. Combining all conditions above then gives the two inequalities in (12a). To complete the proof, we first increase B so that the JSCC rate $B/(B+1)$ is close to one. Fixing this choice of B, we next make n sufficiently large to ensure that all joint typicality requirements behind Claims 5-7 (and similar claims for terminal 2) are satisfied. As now we have $\lim_{n\to\infty}\Pr(\mathcal{E})$=0 (provided that all conditions hold) and ϵ is arbitrary, the distortion pair (D_1, D_2) is achievable. ■

B. Auxiliary Result for Special Case (ii) of Corollary 1

By symmetry, we only show that $I(\tilde{S}_1;\tilde{U}_1|\tilde{S}_2,\tilde{U}_2) < I(\tilde{U}_1;Y_2|\tilde{S}_2,\tilde{U}_2)$ reduces to $R^{(1)}(D_1) < I(X_1;Y_2|X_2)$. First, observe that

$$\begin{aligned}
&I(\tilde{S}_1;\tilde{U}_1|\tilde{S}_2,\tilde{U}_2)\\
&= I(\tilde{S}_1; V_1', \hat{S}_1'|\tilde{S}_2, V_2', \hat{S}_2')\\
&= \underbrace{I(\tilde{S}_1; V_1'|\tilde{S}_2, V_2', \hat{S}_2')}_{=0} + I(\tilde{S}_1; \hat{S}_1'|\tilde{S}_2, V_2', \hat{S}_2', V_1')\\
&= H(\hat{S}_1'|\tilde{S}_2, V_2', \hat{S}_2', V_1') - H(\hat{S}_1'|\tilde{S}_2, V_2', \hat{S}_2', V_1', \tilde{S}_1)\\
&= H(\hat{S}_1') - H(\hat{S}_1'|\tilde{S}_1) \qquad (34)\\
&= I(\tilde{S}_1;\hat{S}_1')
\end{aligned}$$

$$\mathcal{E}_1^{(1)} \triangleq \{(\boldsymbol{S}_1^{(1)}, \boldsymbol{S}_2^{(1)}, \boldsymbol{U}_1^{(1)}(M_1^{(1)}), \boldsymbol{U}_2^{(1)}(M_2^{(1)}), \tilde{\boldsymbol{S}}_1^{(1)}, \tilde{\boldsymbol{S}}_2^{(1)}, \tilde{\boldsymbol{U}}_1^{(1)}, \tilde{\boldsymbol{U}}_2^{(1)}, \tilde{\boldsymbol{W}}_1^{(1)}, \tilde{\boldsymbol{W}}_2^{(1)}, \boldsymbol{X}_1^{(b)}, \boldsymbol{X}_2^{(b)}, \boldsymbol{Y}_1^{(b)}, \boldsymbol{Y}_2^{(b)}) \notin T_\epsilon^{(n)}\}. \tag{29a}$$

$$\begin{aligned}
\mathcal{E}_1^{(B+1)} \triangleq\ &\{(\boldsymbol{S}_1^{(B+1)}, \boldsymbol{S}_2^{(B+1)}, \boldsymbol{U}_1^{(B+1)}, \boldsymbol{U}_2^{(B+1)}, \tilde{\boldsymbol{S}}_1^{(B+1)}, \tilde{\boldsymbol{S}}_2^{(B+1)}, \tilde{\boldsymbol{U}}_1^{(B+1)}(\hat{M}_1^{(B)}), \tilde{\boldsymbol{U}}_2^{(B+1)}(M_2^{(B)}),\\
&\tilde{\boldsymbol{W}}_1^{(B+1)}, \tilde{\boldsymbol{W}}_2^{(B+1)}, \boldsymbol{X}_1^{(B+1)}, \boldsymbol{X}_2^{(B+1)}, \boldsymbol{Y}_1^{(B+1)}, \boldsymbol{Y}_2^{(B+1)}) \notin T_\epsilon^{(n)}\}. \qquad (29b)
\end{aligned}$$

$$\begin{aligned}
\mathcal{E}_1^{(b)} \triangleq\ &\{(\boldsymbol{S}_1^{(b)}, \boldsymbol{S}_2^{(b)}, \boldsymbol{U}_1^{(b)}(M_1^{(b)}), \boldsymbol{U}_2^{(b)}(M_2^{(b)}), \tilde{\boldsymbol{S}}_1^{(b)}, \tilde{\boldsymbol{S}}_2^{(b)}, \tilde{\boldsymbol{U}}_1^{(b)}(\hat{M}_1^{(b-1)}), \tilde{\boldsymbol{U}}_2^{(b)}(M_2^{(b-1)}),\\
&\tilde{\boldsymbol{W}}_1^{(b)}, \tilde{\boldsymbol{W}}_2^{(b)}, \boldsymbol{X}_1^{(b)}, \boldsymbol{X}_2^{(b)}, \boldsymbol{Y}_1^{(b)}, \boldsymbol{Y}_2^{(b)}) \notin T_\epsilon^{(n)}\}, \text{ for } b = 2,3,\ldots,B. \qquad (29c)
\end{aligned}$$

$$= R^{(1)}(D_1) \tag{35}$$

where (34) holds since $\tilde{S}_1$ and $\tilde{S}_2$ are independent and hence $\hat{S}_1'$ is independent of $(\tilde{S}_2, V_2', \hat{S}_2', V_1')$, and (35) follows since the joint probability distribution $P_{\tilde{S}_1,\hat{S}_1'} = P_{S_1,\hat{S}_1}$ achieves $R^{(1)}(D_1)$.

Moreover, we have that

$$\begin{aligned}
&I(\tilde{U}_1; Y_2|\tilde{S}_2, \tilde{U}_2)\\
&= I(V_1', \hat{S}_1'; Y_2|\tilde{S}_2, V_2', \hat{S}_2')\\
&= I(V_1'; Y_2|\tilde{S}_2, V_2', \hat{S}_2') + I(\hat{S}_1'; Y_2|\tilde{S}_2, V_2', \hat{S}_2', V_1')\\
&= I(X_1; Y_2|\tilde{S}_2, X_2, \hat{S}_2') + I(\hat{S}_1'; Y_2|\tilde{S}_2, X_2, \hat{S}_2', X_1) \quad (36)\\
&= H(Y_2|\tilde{S}_2, X_2, \hat{S}_2') - H(Y_2|\tilde{S}_2, X_2, \hat{S}_2', X_1) \quad (37)\\
&= H(Y_2|X_2) - H(Y_2|X_2, X_1) \quad (38)\\
&= I(X_1; Y_2|X_2)
\end{aligned}$$

where (36) follows since $X_j = V_j'$, (37) holds since given channel inputs X_1 and X_2, the output Y_2 is independent of other variables, and (38) holds due to the Markov chain relationship $(\tilde{S}_2, \hat{S}_2') -\!\circ\!- X_2 -\!\circ\!- Y_2$.

C. Proof of Converse Part in Theorem 5

For $k_1 \leq k_2$, let $S_{j,k_1}^{k_2} \triangleq (S_{j,k_1}, S_{j,k_1+1}, \ldots, S_{j,k_2})$. Given a rate-$K/N$ joint source-channel code that achieves the distortion pair (D_1, D_2), we obtain (26a) by the following derivation:

$$\begin{aligned}
&K \cdot R_{S_1|S_0}(D_1)\\
&\leq K \cdot R_{S_1|S_0}\left(K^{-1}\sum_{k=1}^{K}\mathbb{E}\left[d_1(S_{1,k}, \hat{S}_{1,k})\right]\right) \quad (39)\\
&\leq \sum_{k=1}^{K} R_{S_1|S_0}\left(\mathbb{E}[d_1(S_{1,k}, \hat{S}_{1,k})]\right) \quad (40)\\
&\leq \sum_{k=1}^{K} I(S_{1,k}; \hat{S}_{1,k}|S_{0,k}) \quad (41)\\
&\leq \sum_{k=1}^{K} I(S_{1,k}; S_2^K, Y_2^N|S_{0,k}) \quad (42)\\
&\leq \sum_{k=1}^{K} H(S_{1,k}|S_{0,k}) - H(S_{1,k}|S_0^k, S_2^K, Y_2^N) \quad (43)\\
&\leq \sum_{k=1}^{K} H(S_{1,k}|S_0^K, S_1^{k-1}, S_2^K)\\
&\qquad - H(S_{1,k}|S_0^K, S_1^{k-1}, S_2^K, Y_2^N) \quad (44)\\
&= \sum_{k=1}^{K} I(S_{1,k}; Y_2^N|S_0^K, S_1^{k-1}, S_2^K)\\
&= I(S_1^K; Y_2^N|S_0^K, S_2^K)\\
&= \sum_{n=1}^{N} I(S_1^K; Y_{2,n}|S_0^K, S_2^K, Y_2^{n-1})\\
&\leq \sum_{n=1}^{N} H(Y_{2,n}|X_{2,n})\\
&\qquad - H(Y_{2,n}|S_0^K, S_1^K, S_2^K, Y_2^{n-1}, X_{1,n}, X_{2,n}) \quad (45)\\
&= \sum_{n=1}^{N} H(Y_{2,n}|X_{2,n}) - H(Y_{2,n}|X_{1,n}, X_{2,n}) \quad (46)\\
&= N \cdot \sum_{n=1}^{N} \frac{1}{N} \cdot I(X_{1,n}; Y_{2,n}|X_{2,n})\\
&\leq N \cdot I(X_1; Y_2|X_2), \quad (47)
\end{aligned}$$

where (39) holds since $R_{S_1|S_0}(D_1)$ is non-increasing and the expected distortion of the code is not larger than D_1, (40) and (41) are respectively due to convexity and the definition of conditional RD function, (42) follows from the data-processing inequality, (43) holds since conditioning reduces entropy, (44) holds by the Markov chain relationships $S_{1,k} -\!\circ\!- S_{0,k} -\!\circ\!- (S_0^{k-1}, S_{0,k+1}^K, S_1^{k-1})$ and $S_1^K -\!\circ\!- S_0^K -\!\circ\!- S_2^K$ and since conditioning reduces entropy, (45) holds since $X_{2,n}$ is a function of (Y_2^{n-1}, S_2^K) and since conditioning reduces entropy, (46) follows from the memoryless property of channel, and (47) holds with $P_{X_1,X_2} = N^{-1}\sum_{n=1}^{N} P_{X_{1,n},X_{2,n}}$ since $I(X_{1,n}; Y_{2,n}|X_{2,n})$ is concave in $P_{X_{1,n},X_{2,n}}$. By symmetry, a similar argument shows (26b). ■

REFERENCES

[1] J.-J. Weng, F. Alajaji, and T. Linder, "Lossy transmission of correlated sources over two-way channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2017, pp. 354–358.

[2] J.-J. Weng, F. Alajaji, and T. Linder, "Joint source-channel coding for the transmission of correlated sources over two-way channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1322–1326.

[3] J.-J. Weng, F. Alajaji, and T. Linder, "Adaptive coding for two-way lossy source-channel communication," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 2468–2473.

[4] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, 1961, pp. 611–644.

[5] A. Jamalipour, T. Wada, and T. Yamazato, "A tutorial on multiple access technologies for beyond 3G mobile networks," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 110–117, Feb. 2005.

[6] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[7] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, Nov. 2014.

[8] Y. Yuan *et al.*, "Non-orthogonal transmission technology in LTE evolution," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 68–74, Jul. 2016.

[9] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

[10] L. R. Varshney, "Two way communication over exponential family type channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2795–2799.

[11] A. Chaaban, L. R. Varshney, and M.-S. Alouini, "The capacity of injective semi-deterministic two-way channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 431–435.

[12] D. Seo, A. Chaaban, L. R. Varshney, and M.-S. Alouini, "Classes of full-duplex channels with capacity achieved without adaptation," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4141–4149, Jul. 2020.

[13] J.-J. Weng, L. Song, F. Alajaji, and T. Linder, "Capacity of two-way channels with symmetry properties," *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6290–6313, Oct. 2019.

[14] J.-J. Weng, F. Alajaji, and T. Linder, "Capacity of generalized discrete-memoryless push-to-talk two-way channels," in *Proc. 16th Can. Workshop Inf. Theory (CWIT)*, Jun. 2019, pp. 1–6.

[15] T. S. Han, "A general coding scheme for the two-way channel," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 1, pp. 35–44, Jan. 1984.

[16] J. Schalkwijk, "The binary multiplying channel—A coding scheme that operates beyond Shannon's inner bound region (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 107–110, Jan. 1982.

[17] J. Schalkwijk, "On an extension of an achievable rate region for the binary multiplying channel," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 445–448, May 1983.

[18] G. Kramer, "Directed information for channels with feedback," Ph.D. dissertation, Swiss Federal Inst. Technol. Zürich, Zürich, Switzerland, 1998.

[19] O. Sabag and H. H. Permuter, "An achievable rate region for the two-way channel with common output," in *Proc. 56th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2018, pp. 527–531.

[20] Z. Zhang, T. Berger, and J. Schalkwijk, "New outer bounds to capacity regions of two-way channels," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 3, pp. 383–386, May 1986.

[21] A. H. Kaspi, "Two-way source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 735–740, Nov. 1985.

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

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

[24] A. Maor and N. Merhav, "Two-way successively refined joint source-channel coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1483–1494, Apr. 2006.

[25] M. Vera, L. R. Vega, and P. Piantanida, "Collaborative information bottleneck," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 787–815, Feb. 2019.

[26] H. H. Permuter, Y. Steinberg, and T. Weissman, "Two-way source coding with a helper," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2905–2919, Jun. 2010.

[27] Y.-K. Chia, H. H. Permuter, and T. Weissman, "Cascade, triangular, and two-way source coding with degraded side information at the second user," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 189–206, Jan. 2011.

[28] L. R. Vega, P. Piantanida, and A. O. Hero, "The three-terminal interactive lossy source coding problem," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 532–562, Jan. 2017.

[29] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6180–6195, Sep. 2011.

[30] S. Shin and C. Suh, "Two-way function computation," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 813–834, Feb. 2020.

[31] Z. Cheng and N. Devroye, "Two-way networks: When adaptation is useless," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1793–1813, Mar. 2014.

[32] L. Ong, "Capacity results for two classes of three-way channels," in *Proc. Int. Symp. Commun. Inf. Technol. (ISCIT)*, Oct. 2012, pp. 471–476.

[33] L. Ong, G. Lechner, S. J. Johnson, and C. M. Kellett, "The three-user finite-field multi-way relay channel with correlated sources," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3125–3135, Aug. 2013.

[34] A. Chaaban, H. Maier, A. Sezgin, and R. Mathar, "Three-way channels with multiple unicast sessions: Capacity approximation via network transformation," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7086–7102, Dec. 2016.

[35] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3927–3944, Sep. 2009.

[36] T. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[37] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[38] P. Minero, S. H. Lim, and Y.-H. Kim, "A unified approach to hybrid coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1509–1523, Apr. 2015.

[39] R. M. Gray, "Conditional rate-distortion theory," Stanford Univ. Electron. Lab, Stanford, CA, USA, Tech. Rep. 6502-2, 1972.

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

[41] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[43] S.-H. Lee and S.-Y. Chung, "A unified random coding bound," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6779–6802, Oct. 2018.

[44] N. Deligiannis, A. Sechelea, A. Munteanu, and S. Cheng, "The no-rate-loss property of Wyner-Ziv coding in the Z-channel correlation case," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1675–1678, Oct. 2014.

[45] L. Song, F. Alajaji, and T. Linder, "Adaptation is useless for two discrete additive-noise two-way channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1854–1858.

[46] G. Dueck, "The capacity region of the two-way channel can exceed the inner bound," *Inf. Control*, vol. 40, no. 3, pp. 258–266, Mar. 1979.

Jian-Jia Weng (Member, IEEE) received the B.S. degree in electrical engineering from I-Shou University (ISU), Kaohsiung, Taiwan, in 2004, the M.S. degree in communication engineering from National Central University (NCU), Taoyuan, Taiwan, in 2006, and the Ph.D. degree in mathematics and statistics from Queen's University, Kingston, Canada, in 2020. From 2006 to 2014, he was a Ph.D. Student with the Institute of Communications Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan. From July 2012 to May 2013, he was also a Visiting Scholar with the Information Processing Group, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. In 2021, he held a post-doctoral appointment at the Department of Mathematics and Statistics, Queen's University. He is currently an Assistant Professor with the Department of Electrical Engineering, National Taiwan Ocean University (NTOU), Keelung, Taiwan. His research interests include coding theory and information theory.

Fady Alajaji (Senior Member, IEEE) received the B.E. degree (Hons.) in electrical engineering from the American University of Beirut, Lebanon, in 1988, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1990 and 1994, respectively.

In 1994, he held a post-doctoral appointment at the Institute for Systems Research, University of Maryland. In 1995, he joined the Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada, where he is currently a Professor of mathematics and engineering. Since 1997, he has also been cross-appointed with the Department of Electrical and Computer Engineering, Queen's University, where he served as the Acting Head for the Department of Mathematics and Statistics, from 2013 to 2014. From 2003 to 2008 and 2018 to 2019, he served as the Chair for the Queen's Mathematics and Engineering Program. His research interests include information theory, digital communications, error control coding, joint source-channel coding, data compression, network epidemics, and machine learning.

Dr. Alajaji served as an organizer and technical program committee member of several international conferences and workshops. He received the Premier's Research Excellence Award from the Province of Ontario. He is an Associate Editor of Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY. He served as an Area Editor of Source-Channel Coding and Signal Processing (2008–2015) and as an Editor for Source and Source-Channel Coding (2003–2012) for the IEEE TRANSACTIONS ON COMMUNICATIONS.

Tamás Linder (Fellow, IEEE) received the M.S. degree in electrical engineering from the Technical University of Budapest, Hungary, in 1988, and the Ph.D. degree in electrical engineering from the Hungarian Academy of Sciences in 1992.

He was a Post-Doctoral Researcher at the University of Hawaii in 1992 and a Visiting Fulbright Scholar at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, from 1993 to 1994. From 1994 to 1998, he was a Faculty Member of the Department of Computer Science and Information Theory, Technical University of Budapest. From 1996 to 1998, he was also a Visiting Research Scholar with the Department of Electrical and Computer Engineering, University of California at San Diego. In 1998, he joined Queen's University, where he is currently a Professor of mathematics and engineering with the Department of Mathematics and Statistics. His research interests include communications and information theory, source coding and vector quantization, machine learning, and statistical pattern recognition.

Dr. Linder received the Premier's Research Excellence Award of the Province of Ontario in 2002 and the Chancellor's Research Award of Queen's University in 2003. He was an Associate Editor for source coding of the IEEE TRANSACTIONS ON INFORMATION THEORY from 2003 to 2004.