A New Data Processing Inequality and Its Applications in Distributed Source and Channel Coding

Wei Kang, *Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

***Abstract*—In the distributed coding of correlated sources, the problem of characterizing the joint probability distribution of a pair of random variables satisfying an n-letter Markov chain arises. The exact solution of this problem is intractable. In this paper, we seek a single-letter necessary condition for this n-letter Markov chain. To this end, we propose a new data processing inequality on a new measure of correlation through a spectral method. Based on this new data processing inequality, we provide a single-letter necessary condition for the required joint probability distribution. We apply our results to two specific examples involving the distributed coding of correlated sources: multiple-access channel with correlated sources and multiterminal rate-distortion region, and propose new necessary conditions for these two problems.**

***Index Terms*—Correlated sources, data processing inequality, multiterminal rate-distortion region, multiple-access channel.**

I. Problem Formulation

In this paper, we consider a pair of correlated discrete source sequences with length n, $(U^n, V^n) = \{(U_1, V_1), \ldots, (U_n, V_n)\}$, which are independent and identically distributed (i.i.d.) in time, i.e.,

$$p(u^n, v^n) = \prod_{i=1}^{n} p(u_i, v_i) \tag{1}$$

and

$$p(u_i, v_i) = p(u, v), \qquad i = 1, \ldots, n \tag{2}$$

where the joint distribution $p(u,v)$ is defined on the alphabet $\mathcal{U} \times \mathcal{V}$. Let (X_1, X_2) be two random variables defined on the alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ such that (X_1, X_2, U^n, V^n) satisfies

$$p(x_1, x_2, u^n, v^n) = p(u^n, v^n) p(x_1|u^n) p(x_2|v^n) \tag{3}$$

or equivalently[1]

$$X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2.$$

This Markov chain appears in some problems involving the distributed coding of correlated sources. For example, in distributed rate-distortion problem [4]–[6], (X_1, X_2) is used to reconstruct, $(\hat{U}^n, \hat{V}^n)$, an estimate of the sources (U^n, V^n), and in the problem of multiple-access channel with correlated sources [7], [8], (X_1, X_2) is sent though a multiple-access channel in one channel use. In this paper, we study the properties of the above Markov chain, which will be applicable to these specific problems.

The study of the converse proofs of (or the necessary conditions for) the above specific problems raises the following question. We know that the correlation between (X_1, X_2) is limited, if Markov chain $X_1 \longrightarrow U \longrightarrow V \longrightarrow X_2$ is to be satisfied. With the help of more letters of the sources, i.e., $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$ with n larger than 1, the correlation between (X_1, X_2) may increase. The question here is how correlated (X_1, X_2) can be, when n increases. More specifically, can they be arbitrarily correlated? To answer this question, we need to determine the set of all "valid" joint probability distributions $p(x_1, x_2)$, if $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$ is to be satisfied for some n, i.e., for given source pair (U, V), we need to determine the following set:[2]

$$\mathcal{S}_{X_1X_2} \triangleq \left\{ p(x_1, x_2) : \begin{array}{l} (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2 \\ \exists n \in \mathbb{N}^+, p(x_1|u^n), p(x_2|v^n), \\ \text{s.t. } p(x_1, x_2) \\ = \sum_{u^n, v^n} p(x_1|u^n) \\ p(u^n, v^n) p(x_2|v^n) \end{array} \right\} \tag{4}$$

with $p(u^n, v^n)$ satisfying (1) and (2).

We note that it is practically impossible to exhaust the elements in the set $\mathcal{S}_{X_1X_2}$ by searching over all conditional distribution pairs $(p(x_1|u^n), p(x_2|v^n))$ for all possible positive integer n. In other words, determining the set of all possible probability distributions $p(x_1, x_2)$ satisfying the n-letter Markov chain $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$, i.e., the set $\mathcal{S}_{X_1X_2}$, seems computationally intractable. To avoid this problem, we seek a necessary condition for the n-letter Markov chain $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$. The resulting set, say $\mathcal{S}'_{X_1X_2}$, characterized by this computable constraints, will contain the target set $\mathcal{S}_{X_1X_2}$.

[1] $X_1 = f_1(U^n)$ and $X_2 = f_2(V^n)$ are degenerate cases.

[2] We are also interested in determining the set of all "valid" probability distributions $p(x_1, x_2, u_1, v_1)$, if this Markov chain constraint is to be satisfied.

Manuscript received November 04, 2006; revised July 26, 2008, December 30, 2009, and July 14, 2010; accepted August 17, 2010. Date of current version December 27, 2010. This work was supported by the National Science Foundation under Grants CCR 03-11311, CCF 04-47613, CCF 05-14846, CNS 07-16311, and CCF 07-29127. A portion of the work of W. Kang was supported by the National Natural Science Foundation of China under Grants 60902008 and 60972027. The material in this paper was presented at the Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, October 2005, at the Conference on Information Sciences and Systems (CISS), Princeton, NJ, March 2006, and at the IEEE International Symposium on Information Theory (ISIT), Seattle, WA, July 2006.

W. Kang was with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742 USA. He is now with the School of Information Science and Engineering, Southeast University, Nanjing, Jiangsu 210096, China (e-mail: wkang@seu.edu.cn).

S. Ulukus is with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ulukus@umd.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2010.2090211

The most intuitive necessary condition for a Markov chain is the data processing inequality [9, p. 32], i.e., if $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$, then

$$I(X_1;X_2) \leq I(U^n;V^n) = nI(U;V). \tag{5}$$

Since $I(U^n;V^n)$ increases linearly with n, the constraint in (5) will be loose when n is sufficiently large. Although the data processing inequality in its usual form does not prove useful in this problem, we will still use the basic methodology of employing a data processing inequality to find a necessary condition for the n-letter Markov chain under consideration. For this, we will introduce a new measure of correlation, and develop a new data processing inequality based on this new measure of correlation.

Spectral method has been instrumental in the study of some properties of pairs of correlated random variables, especially, those of i.i.d. sequences of pairs of correlated random variables, e.g., common information in [10] and isomorphism in [11]. In this paper, we use spectral method to introduce a new data processing inequality, which provides a necessary condition for the joint distributions satisfying the n-letter Markov chain.

The rest of this paper is organized as follows. In Section II, we introduce new measures of correlation and construct a new data processing inequality on these measures. We then apply the new data processing inequality to two specific examples of distributed coding of correlated sources: multiple-access channel with correlated sources in Section III, and multiterminal rate-distortion problem in Section IV. We end with conclusions.

II. Main Results

A. Some Preliminaries

In this section, we provide some basic results which will be used in our later development. The concepts used here are originally introduced by Witsenhausen in [10] in the context of operator theory. Here, we focus on the finite alphabet case, and derive our results in matrix form.

We first introduce our matrix notation for probability distributions. For a pair of discrete random variables X and Y, which take values in $\mathcal{X}$ and $\mathcal{Y}$, respectively, the $|\mathcal{X}| \times |\mathcal{Y}|$ joint probability distribution matrix P_{XY} is defined as

$$P_{XY}(i,j) \triangleq Pr(X=x_i, Y=y_j) \tag{6}$$

where $P_{XY}(i,j)$ denotes the (i,j)-th element of the matrix P_{XY}. The marginal distribution matrix of a random variable X, P_X, is defined as a diagonal matrix with

$$P_X(i,i) \triangleq Pr(X=x_i) \tag{7}$$

and the vector-form marginal distribution, p_X, is defined as[3]

$$p_X(i) \triangleq Pr(X=x_i) \tag{8}$$

or equivalently $p_X = P_X\mathbf{e}$, where $\mathbf{e}$ is the vector of all ones. p_X can also be defined as $p_X \triangleq P_{XY}$ for some degenerate random variable Y whose alphabet size $|\mathcal{Y}|$ is equal to one. For convenience, we define

$$p_X^{\frac{1}{2}} \triangleq P_X^{\frac{1}{2}}\mathbf{e}. \tag{9}$$

[3]In this paper, we only consider the case where p_X is a positive vector.

For conditional distributions, we define matrix $P_{XY|z}$ as

$$P_{XY|z}(i,j) \triangleq Pr(X=x_i, Y=y_j|Z=z). \tag{10}$$

The vector-form conditional distribution $p_{X|z}$ is defined as

$$p_{X|z}(i) \triangleq Pr(X=x_i|Z=z) \tag{11}$$

or equivalently, $p_{X|z} \triangleq P_{XY|z}$ for some degenerate random variable Y whose alphabet size $|\mathcal{Y}|$ is equal to one.

We define a new matrix, $\tilde{P}_{XY}$, which will play an important role in the rest of the paper, as

$$\tilde{P}_{XY} \triangleq P_X^{-\frac{1}{2}} P_{XY} P_Y^{-\frac{1}{2}}. \tag{12}$$

Since $p_X \triangleq P_{XY}$ for some degenerate random variable Y whose alphabet size $|\mathcal{Y}|$ is equal to one, we define

$$\tilde{p}_X = P_X^{-\frac{1}{2}} P_{XY} P_Y^{-\frac{1}{2}} = P_X^{-\frac{1}{2}} p_X = p_X^{\frac{1}{2}}. \tag{13}$$

The counterparts for conditional distributions, $\tilde{P}_{XY|z}$ and $\tilde{p}_{X|y}$, can be defined similarly.

A valid joint distribution matrix, P_{XY}, is a matrix whose entries are nonnegative and sum to 1. Due to this constraint, not every matrix will qualify as a $\tilde{P}_{XY}$ corresponding to a joint distribution matrix as defined in (12). A necessary and sufficient condition for $\tilde{P}_{XY}$ to correspond to a joint distribution matrix is given in Theorem 1 below, which identifies the spectral properties of $\tilde{P}_{XY}$. Before stating the theorem, we provide a lemma and a definition regarding stochastic matrices, which will be used in the proof of the theorem.

Definition 1 [12, p. 48]: A square matrix T of order n is called (row) stochastic if

$$T(i,j) \geq 0, \qquad i,j = 1,\ldots,n \tag{14}$$

$$\sum_{j=1}^{n} T(i,j) = 1, \qquad i = 1,\ldots,n. \tag{15}$$

Lemma 1 [12, p. 49]: The spectral radius, which is defined as the maximum of the absolute values of the eigenvalues of a matrix, of a stochastic matrix is 1. A nonnegative matrix T is stochastic if and only if $\mathbf{e}$ is an eigenvector of T corresponding to the eigenvalue 1.

Theorem 1: Assume a pair of given marginal distributions P_X and P_Y. A nonnegative matrix P is a joint distribution matrix with marginal distributions P_X and P_Y, i.e., $P\mathbf{e} = p_X \triangleq P_X\mathbf{e}$ and $P^T\mathbf{e} = p_Y \triangleq P_Y\mathbf{e}$, if and only if the singular value decomposition (SVD) of the nonnegative matrix $\tilde{P}$, which is defined as $\tilde{P} \triangleq P_X^{-\frac{1}{2}} P P_Y^{-\frac{1}{2}}$ satisfies

$$\tilde{P} = M\Sigma N^T = p_X^{\frac{1}{2}} \left(p_Y^{\frac{1}{2}}\right)^T + \sum_{i=2}^{l} \sigma_i \boldsymbol{\mu}_i \boldsymbol{\nu}_i^T \tag{16}$$

where $M \triangleq [\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_l]$ and $N \triangleq [\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_l]$ are two matrices such that $M^TM = I$ and $N^TN = I$, $\Sigma \triangleq \text{diag}[\sigma_1,\ldots,\sigma_l]$ and $l = \min(|\mathcal{X}|,|\mathcal{Y}|)$; $\boldsymbol{\mu}_1 = p_X^{\frac{1}{2}}$, $\boldsymbol{\nu}_1 = p_Y^{\frac{1}{2}}$, and $\sigma_1 = 1 \geq \sigma_2 \geq \cdots \geq \sigma_l \geq 0$. That is, all of the singular values of $\tilde{P}$ are between 0 and 1, the largest singular value of $\tilde{P}$ is 1, and the corresponding left and right singular vectors are $p_X^{\frac{1}{2}}$ and $p_Y^{\frac{1}{2}}$.

Proof: We begin with the "if" part. We want to show that for any nonnegative matrix P where the corresponding $\tilde{P} \triangleq P_X^{-\frac{1}{2}} P P_Y^{-\frac{1}{2}}$ satisfies (16), P is a joint distribution matrix with marginal distributions p_X and p_Y. Let $\tilde{P}$ satisfy (16), then

$$\begin{aligned} P\mathbf{e} &= P_X^{\frac{1}{2}} \tilde{P} P_Y^{\frac{1}{2}} \mathbf{e} \\ &= P_X^{\frac{1}{2}} \left(p_X^{\frac{1}{2}} (p_Y^{\frac{1}{2}})^T + \sum_{i=2}^{l} \sigma_i \boldsymbol{\mu}_i \boldsymbol{\nu}_i^T \right) p_Y^{\frac{1}{2}} \\ &= P_X^{\frac{1}{2}} p_X^{\frac{1}{2}} (p_Y^{\frac{1}{2}})^T p_Y^{\frac{1}{2}} + P_X^{\frac{1}{2}} \sum_{i=2}^{l} \sigma_i \boldsymbol{\mu}_i \boldsymbol{\nu}_i^T \boldsymbol{\nu}_1 \\ &= p_X. \end{aligned} \tag{17}$$

Similarly, $P^T\mathbf{e} = p_Y$. Thus, the nonnegative matrix P is a joint distribution matrix with marginal distributions p_X and p_Y.

Now we turn to the "only if" part. We want to show that for any joint distribution matrix P with marginal distributions p_X and p_Y, (16) should be satisfied. We consider a joint distribution P with marginal distributions p_X and p_Y. We need to show that the singular values of $\tilde{P}$ lie in $[0, 1]$, the largest singular value is equal to 1, and $p_X^{\frac{1}{2}}$ and $p_Y^{\frac{1}{2}}$, respectively, are the left and right singular vectors corresponding to the singular value 1. To this end, we first construct a Markov chain $X \longrightarrow Y \longrightarrow Z$ with $P_{XY} = P_{ZY} = P$ (this construction comes from [10]). Note that this also implies $P_X = P_Z$, $\tilde{P}_{XY} = \tilde{P}_{ZY} = \tilde{P}$, and $P_{X|Y} = P_{Z|Y}$. The special structure of the constructed Markov chain provides the following:

$$\begin{aligned} P_{X|Z} &= P_{X|Y} P_{Y|Z} \\ &= P_{X|Y} P_{Y|X} \\ &= P P_Y^{-1} P^T P_X^{-1} \\ &= P_X^{\frac{1}{2}} (P_X^{-\frac{1}{2}} P P_Y^{-\frac{1}{2}})(P_Y^{-\frac{1}{2}} P^T P_X^{-\frac{1}{2}}) P_X^{-\frac{1}{2}} \\ &= P_X^{\frac{1}{2}} \tilde{P} \tilde{P}^T P_X^{-\frac{1}{2}} \end{aligned} \tag{18}$$

which implies that the matrix $P_{X|Z}$ is similar to the matrix $\tilde{P}\tilde{P}^T$ [13, p. 44]. Therefore, all the eigenvalues of $P_{X|Z}$ are the eigenvalues of $\tilde{P}\tilde{P}^T$ as well, and if $\boldsymbol{\nu}$ is a left eigenvector of $P_{X|Z}$ corresponding to an eigenvalue σ, then $P_X^{\frac{1}{2}}\boldsymbol{\nu}$ is a left eigenvector of $\tilde{P}\tilde{P}^T$ corresponding to the same eigenvalue.

We note that $P_{X|Z}^T$ is a stochastic matrix, therefore, from Lemma 1, $\mathbf{e}$ is a left eigenvector of $P_{X|Z}$ corresponding the eigenvalue 1, which is equal to the spectral radius of $P_{X|Z}$. Since $P_{X|Z}$ is similar to $\tilde{P}\tilde{P}^T$, we have that $p_X^{\frac{1}{2}}$ is a left eigenvector of $\tilde{P}\tilde{P}^T$ with eigenvalue 1, and all the eigenvalues of $\tilde{P}\tilde{P}^T$ lie in $[-1, 1]$. In addition, $\tilde{P}\tilde{P}^T$ is a symmetric positive semidefinite matrix, which implies that the eigenvalues of $\tilde{P}\tilde{P}^T$ are real and nonnegative. Since the eigenvalues of $\tilde{P}\tilde{P}^T$ are nonnegative, and the largest eigenvalue is equal to 1, we conclude that all of the eigenvalues of $\tilde{P}\tilde{P}^T$ lie in the interval $[0, 1]$.

The singular values of $\tilde{P}$ are the square roots of the eigenvalues of $\tilde{P}\tilde{P}^T$, and the left singular vectors of $\tilde{P}$ are the eigenvectors of $\tilde{P}\tilde{P}^T$. Thus, the singular values of $\tilde{P}$ lie in $[0, 1]$, the largest singular value is equal to 1, and $p_X^{\frac{1}{2}}$ is a left singular vector corresponding to the singular value 1. The corresponding right singular vector is

$$\begin{aligned} \boldsymbol{\nu}_1^T &= \boldsymbol{\mu}_1^T \tilde{P} = \left(p_X^{\frac{1}{2}}\right)^T P_X^{-\frac{1}{2}} P P_Y^{-\frac{1}{2}} \\ &= \mathbf{e}^T P P_Y^{-\frac{1}{2}} = p_Y^T P_Y^{-\frac{1}{2}} = \left(p_Y^{\frac{1}{2}}\right)^T \end{aligned} \tag{19}$$

which concludes the proof. ■

This theorem implies that there is a one-to-one mapping between all joint distribution matrices P and all nonnegative matrices $\tilde{P}$ satisfying (16). It is easy to see from (12) that there is a corresponding $\tilde{P}$ for every P. Conversely, any given nonnegative matrix $\tilde{P}$ satisfying (16) gives a unique pair of marginal distributions (P_X, P_Y), which is specified by the left and right positive singular vectors corresponding to its largest singular value[4]. Then, from (12), using $\tilde{P}$ and (P_X, P_Y) given by its singular vectors, we obtain a corresponding P as

$$P = P_X^{\frac{1}{2}} \tilde{P} P_Y^{\frac{1}{2}}. \tag{20}$$

Because of this one-to-one relationship, exploring all possible joint distribution matrices P is equivalent to exploring all possible nonnegative matrices $\tilde{P}$ satisfying (16).

Here, $\sigma_2, \ldots, \sigma_l$ can be viewed as a group of quantities, which measures the correlation between random variables X and Y. We note that when $\sigma_2 = \cdots = \sigma_l = 1$, X and Y are fully correlated, and, when $\sigma_2 = \cdots = \sigma_l = 0$, X and Y are independent. In all the cases between these two extremes, X and Y are arbitrarily correlated. Moreover, Witsenhausen showed that X and Y have a common data if and only if $\sigma_2 = 1$ [10]. In the next section, we will propose a new data processing inequality with respect to these new measures of correlation, $\sigma_2, \ldots, \sigma_l$. By utilizing this new data processing inequality, we will provide a necessary condition for the n-letter Markov chain $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$.

B. A New Data Processing Inequality

In this section, first, we introduce a new data processing inequality in the following theorem. Here, we provide a lemma that will be used in the proof of the theorem.

Lemma 2 [14, p. 178]: For matrices A and B

$$\sigma_i(AB) \leq \sigma_i(A)\sigma_1(B) \tag{21}$$

where $\sigma_i(\cdot)$ denotes the ith largest singular value of a matrix.

Theorem 2: If $X \longrightarrow Y \longrightarrow Z$, then

$$\sigma_i(\tilde{P}_{XZ}) \leq \sigma_i(\tilde{P}_{XY})\sigma_2(\tilde{P}_{YZ}) \leq \sigma_i(\tilde{P}_{XY}) \tag{22}$$

where $i = 2, \ldots, \mathrm{rank}(\tilde{P}_{XZ})$.

Proof: From the structure of the Markov chain, and from the definition of $\tilde{P}_{XY}$ in (12), we have

$$\begin{aligned} \tilde{P}_{XZ} &= P_X^{-\frac{1}{2}} P_{XZ} P_Z^{-\frac{1}{2}} \\ &= P_X^{-\frac{1}{2}} P_{XY} P_Y^{-\frac{1}{2}} P_Y^{-\frac{1}{2}} P_{YZ} P_Z^{-\frac{1}{2}} = \tilde{P}_{XY}\tilde{P}_{YZ}. \end{aligned} \tag{23}$$

[4]We observe that there may exist multiple singular values equal to 1, but $\boldsymbol{\mu}_1$ and $\boldsymbol{\nu}_1$ are the only positive singular vectors, because singular vectors are orthonormal.

Using (16) for $\tilde{P}_{XZ}$, we obtain

$$\tilde{P}_{XZ} = p_X^{\frac{1}{2}}(p_Z^{\frac{1}{2}})^T + \sum_{i=2}^{l} \sigma_i(\tilde{P}_{XZ})\boldsymbol{\mu}_i(\tilde{P}_{XZ})\boldsymbol{\nu}_i(\tilde{P}_{XZ})^T \quad (24)$$

and applying (16) to $\tilde{P}_{XY}$ and $\tilde{P}_{YZ}$ yields

$$\begin{aligned}\tilde{P}_{XY}\tilde{P}_{YZ} &= \left(p_X^{\frac{1}{2}}(p_Y^{\frac{1}{2}})^T + \sum_{i=2}^{l} \sigma_i(\tilde{P}_{XY})\boldsymbol{\mu}_i(\tilde{P}_{XY})\boldsymbol{\nu}_i(\tilde{P}_{XY})^T\right) \\ &\quad \times \left(p_Y^{\frac{1}{2}}(p_Z^{\frac{1}{2}})^T + \sum_{i=2}^{l} \sigma_i(\tilde{P}_{YZ})\boldsymbol{\mu}_i(\tilde{P}_{YZ})\boldsymbol{\nu}_i(\tilde{P}_{YZ})^T\right) \\ &= p_X^{\frac{1}{2}}(p_Z^{\frac{1}{2}})^T + \left(\sum_{i=2}^{l} \sigma_i(\tilde{P}_{XY})\boldsymbol{\mu}_i(\tilde{P}_{XY})\boldsymbol{\nu}_i(\tilde{P}_{XY})^T\right) \\ &\quad \times \left(\sum_{i=2}^{l} \sigma_i(\tilde{P}_{YZ})\boldsymbol{\mu}_i(\tilde{P}_{YZ})\boldsymbol{\nu}_i(\tilde{P}_{YZ})^T\right) \quad (25)\end{aligned}$$

where the two cross-terms vanish because $p_Y^{\frac{1}{2}}$ plays the roles of both $\boldsymbol{\nu}_1(\tilde{P}_{XY})$ and $\boldsymbol{\mu}_1(\tilde{P}_{YZ})$, and therefore, $p_Y^{\frac{1}{2}}$ is orthogonal to both $\boldsymbol{\nu}_i(\tilde{P}_{XY})$ and $\boldsymbol{\mu}_j(\tilde{P}_{YZ})$, for all $i,\ j \neq 1$. Using (23) and equating (24) and (25), we obtain

$$\begin{aligned}&\sum_{i=2}^{l} \sigma_i(\tilde{P}_{XZ})\boldsymbol{\mu}_i(\tilde{P}_{XZ})\boldsymbol{\nu}_i(\tilde{P}_{XZ})^T \\ &\quad = \left(\sum_{i=2}^{l} \sigma_i(\tilde{P}_{XY})\boldsymbol{\mu}_i(\tilde{P}_{XY})\boldsymbol{\nu}_i(\tilde{P}_{XY})^T\right) \\ &\qquad \times \left(\sum_{i=2}^{l} \sigma_i(\tilde{P}_{YZ})\boldsymbol{\mu}_i(\tilde{P}_{YZ})\boldsymbol{\nu}_i(\tilde{P}_{YZ})^T\right). \quad (26)\end{aligned}$$

The proof is completed by applying Lemma 2 to (26) and also by noting that $\sigma_2(\tilde{P}_{YZ}) \leq 1$ from Theorem 1. ■

Theorem 2 is a new data processing inequality in the sense that the processing from Y to Z reduces the correlation measure σ_i, i.e., the correlation between X and Z, $\sigma_i(\tilde{P}_{XZ})$, is less than or equal to the correlation measure between X and Y, $\sigma_i(\tilde{P}_{XY})$. We note that this theorem is similar to the data processing inequality in [9, p. 32] except instead of mutual information, we use $\sigma_i(\tilde{P}_{XY})$ as the correlation measure. In the sequel, we will show that this new data processing inequality helps us develop a necessary condition for the n-letter Markov chain while the data processing inequality in its usual form [9, p. 32] is not useful in this context.

C. A Necessary Condition for the n-Letter Markov Chain

Now, we switch our attention to i.i.d. sequences of correlated sources. Let (U^n, V^n) be a pair of i.i.d. (in time) sequences, where each letter of these sequences satisfies a joint distribution P_{UV}. Thus, the joint distribution of the sequences is $P_{U^nV^n} = P_{UV}^{\otimes n}$, where $A^{\otimes 1} \triangleq A$, $A^{\otimes k} \triangleq A \otimes A^{\otimes (k-1)}$, and $\otimes$ denotes the Kronecker product of matrices [13].

From (12), we know that

$$P_{UV} = P_U^{\frac{1}{2}}\tilde{P}_{UV}P_V^{\frac{1}{2}}. \quad (27)$$

Then

$$\begin{aligned}P_{U^nV^n} &= P_{UV}^{\otimes n} = \left(P_U^{\frac{1}{2}}\tilde{P}_{UV}P_V^{\frac{1}{2}}\right)^{\otimes n} \\ &= \left(P_U^{\frac{1}{2}}\right)^{\otimes n}\tilde{P}_{UV}^{\otimes n}\left(P_V^{\frac{1}{2}}\right)^{\otimes n}. \quad (28)\end{aligned}$$

We also have $P_{U^n} = P_U^{\otimes n}$ and $P_{V^n} = P_V^{\otimes n}$. Thus

$$\begin{aligned}\tilde{P}_{U^nV^n} &\triangleq P_{U^n}^{-\frac{1}{2}}P_{U^nV^n}P_{V^n}^{-\frac{1}{2}} \\ &= \left(P_U^{-\frac{1}{2}}\right)^{\otimes n}\left(P_U^{\frac{1}{2}}\right)^{\otimes n} \\ &\quad \times \tilde{P}_{UV}^{\otimes n}\left(P_V^{\frac{1}{2}}\right)^{\otimes n}\left(P_V^{-\frac{1}{2}}\right)^{\otimes n} \\ &= \tilde{P}_{UV}^{\otimes n}. \quad (29)\end{aligned}$$

Now, applying SVD to $\tilde{P}_{U^nV^n}$, we have

$$\tilde{P}_{U^nV^n} = M_n\Sigma_nN_n^T = \tilde{P}_{UV}^{\otimes n} = M^{\otimes n}\Sigma^{\otimes n}\left(N^{\otimes n}\right)^T. \quad (30)$$

From the uniqueness of the SVD, we know that $M_n = M^{\otimes n}$, $\Sigma_n = \Sigma^{\otimes n}$ and $N_n = N^{\otimes n}$. Then, the ordered singular values of $\tilde{P}_{U^nV^n}$ are

$$\left\{1, \sigma_2(\tilde{P}_{UV}), \ldots, \sigma_2(\tilde{P}_{UV}), \ldots\right\} \quad (31)$$

where the second through the $n+1$st singular values are all equal to $\sigma_2(\tilde{P}_{UV})$.

From Theorem 2, we know that if $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$, then, for $i \geq 2$

$$\sigma_i(\tilde{P}_{X_1X_2}) \leq \sigma_2(\tilde{P}_{X_1U^n})\sigma_i(\tilde{P}_{U^nV^n})\sigma_2(\tilde{P}_{V^nX_2}). \quad (32)$$

As shown in (31), we have $\sigma_i(\tilde{P}_{U^nV^n}) \leq \sigma_2(\tilde{P}_{UV})$ for $i \geq 2$. Therefore, for $i \geq 2$, we have

$$\sigma_i(\tilde{P}_{X_1X_2}) \leq \sigma_2(\tilde{P}_{X_1U^n})\sigma_2(\tilde{P}_{UV})\sigma_2(\tilde{P}_{V^nX_2}). \quad (33)$$

From Theorem 1, we know that $\sigma_2(\tilde{P}_{X_1U^n}) \leq 1$ and $\sigma_2(\tilde{P}_{V^nX_2}) \leq 1$.

Based on the above discussion, we have the following theorem.

Theorem 3: If $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$, then, we have

$$\sigma_i(\tilde{P}_{X_1X_2}) \leq \sigma_2(\tilde{P}_{UV}), \quad i \geq 2. \quad (34)$$

Theorem 3 provides a necessary condition for the n-letter Markov chain $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$ on the joint probability distribution $p(x_1, x_2)$. The set characterized by this condition is defined as follows:

$$\mathcal{S}'_{X_1X_2} \triangleq \left\{p(x_1,x_2) : \begin{array}{l}(x_1,x_2) \in \mathcal{X}_1 \times \mathcal{X}_2 \\ \sigma_i(\tilde{P}_{X_1X_2}) \leq \sigma_2(\tilde{P}_{UV}),\ \text{for } i \geq 2\end{array}\right\}. \quad (35)$$

From Theorem 3, we have

$$\mathcal{S}_{X_1X_2} \subseteq \mathcal{S}'_{X_1X_2} \quad (36)$$

where $\mathcal{S}_{X_1X_2}$ is defined in (4).

Theorem 3 answers the question we posed in Section I. Our question was whether (X_1, X_2) can be arbitrarily correlated, when we allow n to take any value in $\mathbb{N}^+$. Theorem 3 shows that (X_1, X_2) cannot be arbitrarily correlated, as the correlation measures between (X_1, X_2), $\sigma_i(\tilde{P}_{X_1X_2})$, are upper bounded by, $\sigma_2(\tilde{P}_{UV})$, the second correlation measure of the sources (U, V), no matter what value n takes.

As we mentioned in Section I, the data processing inequality in its usual form [9, p. 32] is not helpful in this problem, while our new data processing inequality, i.e., Theorem 2, provides a necessary condition for this n-letter Markov chain. The main reason for this difference is that while the mutual information, $I(U^n; V^n)$, the correlation measure in the original data processing inequality, increases linearly with n, $\sigma_i(\tilde{P}_{U^nV^n})$, the correlation measure in our new data processing inequality, is bounded as n increases, and therefore, makes the problem more tractable.

Theorem 3 is valid for all discrete random variables. A sharper result in a special binary case can be found in [15].

D. Conditional Probability

Theorem 3 in Section II-C provides a necessary condition for joint probability distributions $p(x_1, x_2)$, which satisfy the Markov chain $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$. In certain specific problems, e.g., multiterminal rate-distortion problem and multiple-access channel with correlated sources, in addition to $p(x_1, x_2)$, we are also interested in the conditional distribution[5] $p(x_1, x_2|u_1, v_1)$. In this section, we will develop a result similar to that in Theorem 3 for conditional distributions.

We wish to determine the set of all possible conditional distributions $p(x_1, x_2|u_1, v_1)$ satisfying $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$, i.e., the following set:

$$\mathcal{S}_{X_1X_2|UV} \triangleq \left\{ \begin{array}{l} p(x_1, x_2|u_1, v_1): \\ (x_1, x_2, u_1, v_1) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{U} \times \mathcal{V} \\ \exists n \in \mathbb{N}^+, p(x_1|u^n), p(x_2|v^n) \\ \text{s.t. } p(x_1, x_2|u_1, v_1) = \\ \frac{\sum_{u_2,\dots,u_n,v_2\dots,v_n} p(x_1|u^n)p(x_2|v^n)p(u^n,v^n)}{p(u_1,v_1)} \end{array} \right\} \tag{37}$$

with $p(u^n, v^n)$ satisfying (1) and (2). Due to the same reason as in the case of $\mathcal{S}_{X_1X_2}$, it is practically impossible to exhaust all the elements in the set $\mathcal{S}_{X_1X_2|UV}$. Thus, we seek a set $\mathcal{S}'_{X_1X_2|UV}$, which contains $\mathcal{S}_{X_1X_2|UV}$ as a subset and has a simple description.

We note that $p(x_1, x_2)$, $p(x_1, x_2|u_1)$ and $p(x_1, x_2|v_1)$ are all functions of $p(x_1, x_2|u_1, v_1)$ for a given $p(u_1, v_1)$, i.e.,

$$p(x_1, x_2) = \sum_{u_1, v_1} p(x_1, x_2|u_1, v_1)p(u_1, v_1) \tag{38}$$

$$p(x_1, x_2|u_1) = \frac{\sum_{v_1} p(x_1, x_2|u_1, v_1)p(u_1, v_1)}{\sum_{v_1} p(u_1, v_1)} \tag{39}$$

$$p(x_1, x_2|v_1) = \frac{\sum_{u_1} p(x_1, x_2|u_1, v_1)p(u_1, v_1)}{\sum_{u_1} p(u_1, v_1)}. \tag{40}$$

[5]The reader may wish to consult Sections III and IV for further motivations to consider conditional probability distributions $p(x_1, x_2|u_1, v_1)$.

Thus, $\sigma_i(\tilde{P}_{X_1X_2})$, $\sigma_i(\tilde{P}_{X_1X_2|u_1})$, $\sigma_i(\tilde{P}_{X_1X_2|v_1})$, as well as $\sigma_i(\tilde{P}_{X_1X_2|u_1v_1})$, where the conditional probability matrix is defined in (10), are all functions of $p(x_1, x_2|u_1, v_1)$ for a given $p(u_1, v_1)$. We have the following theorem to characterize the constraints on $\sigma_i(\tilde{P}_{X_1X_2})$, $\sigma_i(\tilde{P}_{X_1X_2|u_1})$, $\sigma_i(\tilde{P}_{X_1X_2|v_1})$, and $\sigma_i(\tilde{P}_{X_1X_2|u_1v_1})$.

Theorem 4: Let (U^n, V^n) be a pair of i.i.d. sequences of length n, and let the random variables X_1, X_2 satisfy $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$. Assume $\underline{U}$ is an arbitrary subset of $\{U_1, \dots, U_n\}$, i.e.,

$$\underline{U} \triangleq \{U_{i_1}, \dots, U_{i_l}\} \subset \{U_1, \dots, U_n\} \tag{41}$$

and similarly

$$\underline{V} \triangleq \{V_{j_1}, \dots, V_{j_k}\} \subset \{V_1, \dots, V_n\}. \tag{42}$$

Then

$$\sigma_i(\tilde{P}_{X_1X_2|\underline{u}\,\underline{v}}) \leq \sigma_2(\tilde{P}_{UV}), \quad i \geq 2. \tag{43}$$

The proof of this theorem can be found in Appendix I.

We define the set $\mathcal{S}'_{X_1X_2|UV}$ as follows:

$$\mathcal{S}'_{X_1X_2|UV} \triangleq \left\{ \begin{array}{l} p(x_1, x_2|u_1, v_1): \\ (x_1, x_2, u_1, v_1) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{U} \times \mathcal{V} \\ \sigma_i(\tilde{P}_{X_1X_2}) \leq \sigma_2(\tilde{P}_{UV}) \quad i \geq 2 \\ \sigma_i(\tilde{P}_{X_1X_2|u_1}) \leq \sigma_2(\tilde{P}_{UV}) \quad i \geq 2 \\ \sigma_i(\tilde{P}_{X_1X_2|v_1}) \leq \sigma_2(\tilde{P}_{UV}) \quad i \geq 2 \\ \sigma_i(\tilde{P}_{X_1X_2|u_1v_1}) \leq \sigma_2(\tilde{P}_{UV}) \quad i \geq 2 \end{array} \right\}. \tag{44}$$

By applying Theorem 4 on $p(x_1, x_2)$, $p(x_1, x_2|u_1)$, $p(x_1, x_2|v_1)$ and $p(x_1, x_2|u_1, v_1)$, respectively, we obtain

$$\mathcal{S}_{X_1X_2|UV} \subseteq \mathcal{S}'_{X_1X_2|UV}. \tag{45}$$

III. Example I: Multiple-Access Channel With Correlated Sources

The problem of determining the capacity region of the multiple-access channel with correlated sources can be formulated as follows. Consider a pair of i.i.d. correlated sources (U, V) described by the joint probability distribution $p(u, v)$ defined on finite alphabet $\mathcal{U} \times \mathcal{V}$. Assume a discrete, memoryless, multiple-access channel characterized by the transition probability $p(y|x_1, x_2)$ defined on finite alphabet $\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}$. A block code (f_1, f_2, g) is defined as

$$f_1 : \mathcal{U}^n \longmapsto \mathcal{X}_1^n \tag{46}$$
$$f_2 : \mathcal{V}^n \longmapsto \mathcal{X}_2^n \tag{47}$$
$$g : \mathcal{Y}^n \longmapsto (\mathcal{U} \times \mathcal{V})^n \tag{48}$$

that is, the transmitter f_1 maps the source U^n to channel input X_1^n, similarly, the transmitter f_2 maps the source V^n to channel input X_2^n, and the receiver g reconstruct $(\hat{U}^n, \hat{V}^n)$ from Y^n. The probability of error is

$$P_e \triangleq \sum_{(u^n,v^n)\in\mathcal{U}^n\times\mathcal{V}^n} Pr((\hat{u}^n, \hat{v}^n) \neq (u^n, v^n)|(U^n, V^n) = (u^n, v^n)) \, Pr((U^n, V^n) = (u^n, v^n)). \tag{49}$$

The problem of interest is to determine for every $0 < \epsilon < 1$ and every sufficiently large n, whether there exists an n-length block code (f_1, f_2, g) such that $P_e \leq \epsilon$.

A. Existing Results

The multiple-access channel with correlated sources was studied by Cover, El Gamal, and Salehi in [7] (a simpler proof was given in [8]), where an achievable region expressed by single-letter entropies and mutual informations was given as follows.

Theorem 5 [7]: A source (U,V) with joint distribution $p(u,v)$ can be sent with arbitrarily small probability of error over a multiple-access channel characterized by $p(y|x_1,x_2)$, if there exist probability mass functions $p(s)$, $p(x_1|u,s)$, $p(x_2|v,s)$, such that

$$H(U|V) < I(X_1;Y|X_2,V,S) \tag{50}$$
$$H(V|U) < I(X_2;Y|X_1,U,S) \tag{51}$$
$$H(U,V|W) < I(X_1,X_2;Y|W,S) \tag{52}$$
$$H(U,V) < I(X_1,X_2;Y) \tag{53}$$

where

$$p(s,u,v,x_1,x_2,y) = p(s)p(u,v)p(x_1|u,s)p(x_2|v,s)p(y|x_1,x_2) \tag{54}$$

and

$$w = f(u) = g(v) \tag{55}$$

is the common information in the sense of Gacs and Korner (see [10]).

The above region can be simplified if there is no common information between U and V as follows [7]:

$$H(U|V) < I(X_1;Y|X_2,V) \tag{56}$$
$$H(V|U) < I(X_2;Y|X_1,U) \tag{57}$$
$$H(U,V) < I(X_1,X_2;Y) \tag{58}$$

where

$$p(u,v,x_1,x_2,y) = p(u,v)p(x_1|u)p(x_2|v)p(y|x_1,x_2). \tag{59}$$

This achievable region was shown to be suboptimal by Dueck [16].

Cover, El Gamal, and Salehi [7] also provided a capacity result with both achievability and converse in the form of some incomputable n-letter mutual informations. Their result is restated in the following theorem.

Theorem 6 [7]: The correlated sources (U,V) can be communicated reliably over the discrete memoryless multiple-access channel $p(y|x_1,x_2)$ if and only if

$$[H(U|V), H(V|U), H(U,V)] \in \bigcup_{n=1}^{\infty} \mathcal{C}_n \tag{60}$$

where

$$\mathcal{C}_n = \left\{ [R_1,R_2,R_3] : \begin{array}{l} R_1 < \frac{1}{n} I(X_1^n;Y^n|X_2^n,V^n) \\ R_2 < \frac{1}{n} I(X_2^n;Y^n|X_1^n,U^n) \\ R_3 < \frac{1}{n} I(X_1^n,X_2^n;Y^n) \end{array} \right\} \tag{61}$$

for some

$$p(u^n,v^n,x_1^n,x_2^n,y^n) = p(x_1^n|u^n)p(x_2^n|v^n) \prod_{i=1}^{n} p(u_i,v_i) \prod_{i=1}^{n} p(y_i|x_{1i},x_{2i}) \tag{62}$$

i.e., for some X_1^n and X_2^n that satisfy the Markov chain $X_1^n \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2^n$.

Some recent results on the transmission of correlated sources over multiple-access channels can be found in [17].

B. New Outer Bound

We propose a new outer bound for the multiple-access channel with correlated sources as follows.

Theorem 7: If a pair of i.i.d. sources (U,V) with joint distribution $p(u,v)$ can be transmitted reliably through a discrete, memoryless, multiple-access channel characterized by $p(y|x_1,x_2)$, then

$$H(U|V) \leq I(X_1;Y|X_2,V,Q) \tag{63}$$
$$H(V|U) \leq I(X_2;Y|X_1,U,Q) \tag{64}$$
$$H(U,V) \leq I(X_1,X_2;Y|Q) \tag{65}$$

where random variables X_1, X_2, and Q are such that

$$p(x_1,x_2,y,u,v,q) = p(q)p(u,v)p(y|x_1,x_2)p(x_1,x_2|u,v,q) \tag{66}$$

and for every given q,

$$p(x_1,x_2|u,v,Q=q) \in \mathcal{S}_{X_1X_2|UV} \subset \mathcal{S}'_{X_1X_2|UV} \tag{67}$$

with $\mathcal{S}_{X_1X_2|UV}$ defined in (37) and $\mathcal{S}'_{X_1X_2|UV}$ defined in (44). The size of the alphabet of Q satisfies $|\mathcal{Q}| \leq 3$.

The proof of this theorem can be found in Appendix II.

C. Numerical Example

In this section, we give some simple numerical examples to show the improvement our proposed outer bound provides with respect to the cut-set bound [9]. For simplicity, we only consider the sum-rate here, i.e., comparing $H(U,V)$ with certain mutual information terms. Assume a multiple-access channel where the alphabets of X_1, X_2 and Y are all binary, and the channel transition probability matrix $p(y|x_1,x_2)$ is given as

$Y \backslash X_1X_2$	11	10	01	00
1	1	1/2	1/2	0
0	0	1/2	1/2	1

The following is the cut-set bound for the sum-rate, which we provide as a benchmark

$$H(U,V) < R_{\text{out},1} \triangleq \max_{p(x_1,x_2)} I(X_1,X_2;Y) = 1 \tag{68}$$

where the maximization is over all binary bivariate distributions. The maximum is achieved by $P(X_1=1,X_2=1)=P(X_1=0,X_2=0)=1/2$. We note that the cut-set bound does not depend on the source distribution. We specify the single-letter necessary condition we proposed in the above section and obtain the following upper bound on the sum-rate

$$H(U,V) < R_{\mathrm{out},2} \triangleq \max_{p(x_1,x_2):\sigma_2(\tilde{P}_{X_1X_2})\le\sigma_2(\tilde{P}_{UV})} I(X_1,X_2;Y). \quad (69)$$

Note that we are using a weakened version of our outer bound in Theorem 7. Theorem 7 restricts probability distribution $p(x_1,x_2,u,v)$ or equivalently $p(x_1,x_2|u,v)$, by imposing constraints on $\sigma_i(\tilde{P}_{X_1X_2})$, $\sigma_i(\tilde{P}_{X_1X_2|U})$, $\sigma_i(\tilde{P}_{X_1X_2|V})$, and $\sigma_i(\tilde{P}_{X_1X_2|UV})$ via the definition of set $\mathcal{S}'_{X_1X_2|UV}$. Here we impose constraint only on probability distribution $p(x_1,x_2)$, which yields a weaker necessary condition in this specific correlated sources through MAC channels problem.

We also consider the achievable sum-rate proposed in [7]

$$H(U,V) \le R_{\mathrm{in}} \triangleq \max_{X_1\longrightarrow U\longrightarrow V\longrightarrow X_2} I(X_1,X_2;Y). \quad (70)$$

We are considering a joint source-channel coding problem. Thus, the bounds we discuss here only provide an answer to the question whether reliable communication is possible or not, by comparing the joint source entropy $H(U,V)$ with the outer bounds, which are the maximum of the mutual information term $I(X_1,X_2;Y)$ subject to different constraints on the probability distributions. This maximum mutual information is different with different bounds, e.g., it is $R_{\mathrm{out},1}$ in the cut-set bound and $R_{\mathrm{out},2}$ in our bound. If $H(U,V)$ is larger than $R_{\mathrm{out},1}$ or $R_{\mathrm{out},2}$, then we conclude that reliable communication is impossible; while if $H(U,V)$ is less than $R_{\mathrm{out},1}$ and $R_{\mathrm{out},2}$, we cannot draw any conclusions as to whether reliable communication is possible or not. On the other hand, the tightness of our upper bound is measured by the gap between our upper bound $R_{\mathrm{out},2}$ and the inner bound R_{in}.

First, we consider a binary source (U,V) with the following joint distribution $p(u,v)$

$U\backslash V$	1	0
1	1/3	1/6
0	1/6	1/3

In this case

$$R_{\mathrm{in}} = R_{\mathrm{out},2} = \frac{2}{3} < R_{\mathrm{out},1} = 1 < H(U,V) = 1.92. \quad (71)$$

Since $R_{\mathrm{out},1} < H(U,V)$, it is impossible to transmit this source through the given channel reliably. We also note that, for this case, our upper bound coincides with the single-letter achievability expression, which means that our upper bound on sum-rate is tight. We shall emphasize that a tight upper bound in this joint source-channel coding problem does not imply the possibility of the reliable transmission.

Next, we consider a binary source (U,V) with the following joint distribution $p(u,v)$:

$U\backslash V$	1	0
1	0	0.1
0	0.1	0.8

In this case

$$R_{\mathrm{in}} = 0.51 < R_{\mathrm{out},2} = 0.56 < H(U,V) = 0.92 < R_{\mathrm{out},1} = 1. \quad (72)$$

We note that, in this case, the value of $H(U,V)$ falls between $R_{\mathrm{out},1}$ and $R_{\mathrm{out},2}$, which means that the cut-set bound in (68) fails to test whether it is impossible to have reliable transmission, while our upper bound determines conclusively that reliable transmission is impossible.

Finally, we consider a binary source (U,V) with the following joint distribution $p(u,v)$

$U\backslash V$	1	0
1	0	0.85
0	0.1	0.05

In this case

$$R_{\mathrm{in}} = 0.57 < H(U,V) = 0.75 < R_{\mathrm{out},2} = 0.9 < R_{\mathrm{out},1} = 1. \quad (73)$$

Since $H(U,V)$ is larger than R_{in} and smaller than $R_{\mathrm{out},2}$, we cannot conclude whether it is possible (or not) to transmit these sources through the channel reliably.

IV. Example II: Multiterminal Rate-Distortion Region

Ever since the milestone paper of Wyner and Ziv [18] on the rate-distortion function of a single source with side information at the decoder, there has been a significant amount of efforts directed towards solving a generalization of this problem, the so called multiterminal rate-distortion problem. Among all the attempts on this difficult problem, the notable works by Tung [4] and Housewright [5] (see also [6]) provide the inner and outer bounds for the rate-distortion region. A more recent progress on this problem is by Wagner and Anantharam in [19], where a tighter outer bound is given.

The multiterminal rate-distortion problem can be formulated as follows. Consider a pair of discrete memoryless sources (U,V), with joint distribution $p(u,v)$ defined on the finite alphabet $\mathcal{U}\times\mathcal{V}$. The reconstruction of the sources is built on another finite alphabet $\hat{\mathcal{U}}\times\hat{\mathcal{V}}$. The distortion measures are defined as $d_1 : \mathcal{U}\times\hat{\mathcal{U}} \longmapsto \mathbb{R}^+\cup\{0\}$ and $d_2 : \mathcal{V}\times\hat{\mathcal{V}} \longmapsto \mathbb{R}^+\cup\{0\}$. Assume that two distributed encoders are functions $f_1 : \mathcal{U}^n \longmapsto \{1,2,\ldots,M_1\}$ and $f_2 : \mathcal{V}^n \longmapsto \{1,2,\ldots,M_2\}$ and a joint decoder is the function $g : \{1,2,\ldots,M_1\}\times\{1,2,\ldots,M_2\} \longmapsto \hat{\mathcal{U}}^n\times\hat{\mathcal{V}}^n$, where n is a positive integer. A pair of distortion levels $\mathbf{D}\triangleq(D_1,D_2)$ is said to be $\mathbf{R}$-attainable, for some rate pair $\mathbf{R}\triangleq(R_1,R_2)$, if for all $\epsilon>0$ and $\delta>0$, there exist, some positive integer n and a set of distributed encoders and joint decoder (f_1,f_2,g) with rates[6] $(\frac{1}{n}\log_2 M_1, \frac{1}{n}\log_2 M_2) = (R_1+\delta, R_2+\delta)$, such that the distortion between the sources (U^n,V^n) and the decoder output $(\hat{U}^n,\hat{V}^n)$ satisfies $\left(Ed_1(U^n,\hat{V}^n), Ed_2(V^n,\hat{V}^n)\right) \le (D_1+\epsilon, D_2+\epsilon)$ where $d_1(U^n,\hat{U}^n) \triangleq \frac{1}{n}\sum_{i=1}^n d_1(U_i,\hat{U}_i)$ and $d_2(V^n,\hat{V}^n) \triangleq \frac{1}{n}\sum_{i=1}^n d_2(V_i,\hat{V}_i)$. The problem here is to determine, for a fixed $\mathbf{D}$, the set $\mathcal{R}(\mathbf{D})$ of all rate pairs $\mathbf{R}$, for which $\mathbf{D}$ is $\mathbf{R}$-attainable.

[6]By $(A,B)<(C,D)$, we mean both $A<B$ and $C<D$, and $(A,B)\le(C,D)$ is defined in a similar manner.

A. Existing Results

We restate the inner bound provided in [4] and [5] in the following theorem.

Theorem 8 [4], [5]: $\mathcal{R}(\mathbf{D}) \supseteq \mathcal{R}_{\text{in}}(\mathbf{D})$, where $\mathcal{R}_{\text{in}}(\mathbf{D})$ is the set of all $\mathbf{R}$ such that there exists a pair of discrete random variables (X_1, X_2), for which the following three conditions are satisfied:

1) The joint distribution satisfies

$$X_1 \longrightarrow U \longrightarrow V \longrightarrow X_2. \tag{74}$$

2) The rate pair satisfies

$$R_1 \geq I(U,V;X_1|X_2) \tag{75}$$
$$R_2 \geq I(U,V;X_2|X_1) \tag{76}$$
$$R_1 + R_2 \geq I(U,V;X_1,X_2). \tag{77}$$

3) There exists $\big(\hat{U}(X_1,X_2), \hat{V}(X_1,X_2)\big)$ such that $\big(Ed_1(U,\hat{U}), Ed_2(V,\hat{V})\big) \leq \mathbf{D}$.

An outer bound is also given in [4] and [5] as follows.

Theorem 9 [4], [5]: $\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},1}(\mathbf{D})$, where $\mathcal{R}_{\text{out},1}(\mathbf{D})$ is the set of all $\mathbf{R}$ such that there exists a pair of discrete random variables (X_1, X_2), for which the following three conditions are satisfied:

1) The joint distribution satisfies

$$X_1 \longrightarrow U \longrightarrow V \tag{78}$$
$$U \longrightarrow V \longrightarrow X_2. \tag{79}$$

2) The rate pair satisfies

$$R_1 \geq I(U,V;X_1|X_2) \tag{80}$$
$$R_2 \geq I(U,V;X_2|X_1) \tag{81}$$
$$R_1 + R_2 \geq I(U,V;X_1,X_2). \tag{82}$$

3) There exists $\big(\hat{U}(X_1,X_2), \hat{V}(X_1,X_2)\big)$ such that $\big(Ed_1(U,\hat{U}), Ed_2(V,\hat{V})\big) \leq \mathbf{D}$.

A tighter upper bound was recently proposed by Wagner and Anantharam as follows.[7]

Theorem 10 [19]: $\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},2}(\mathbf{D})$, where $\mathcal{R}_{\text{out},2}(\mathbf{D})$ is the set of all $\mathbf{R}$ such that there exists a pair of discrete random variables (X_1, X_2), for which the following three conditions are satisfied:

1) The joint distribution satisfies

$$p(x_1,x_2|u,v) : \exists \text{ random variable } W,$$
$$p(x_1,x_2|u,v) = \sum_w p(w)p(x_1|w,u)p(x_2|w,v). \tag{83}$$

This distribution may be represented by the following Markov chain-like notation:

$$\begin{array}{ccccccc} X_1 & \longrightarrow & U & \longrightarrow & V & \longrightarrow & X_2 \\ & \searrow & & & & \nearrow & \\ & & & W & & & \end{array} \tag{84}$$

2) The rate pair satisfies

$$R_1 \geq I(U,V;X_1|X_2) \tag{85}$$
$$R_2 \geq I(U,V;X_2|X_1) \tag{86}$$
$$R_1 + R_2 \geq I(U,V;X_1,X_2). \tag{87}$$

3) There exists $\big(\hat{U}(X_1,X_2), \hat{V}(X_1,X_2)\big)$ such that $\big(Ed_1(U,\hat{U}), Ed_2(V,\hat{V})\big) \leq \mathbf{D}$.

We note that the above three bounds agree on both the second condition, i.e., the rate constraints in terms of some mutual information expressions, and the third condition, i.e., the reconstruction functions. However, the first condition in these three bounds constraining the underlying probability distributions $p(x_1,x_2|u,v)$ are different. It is easy to see that the Markov chain condition in the inner bound, i.e., $X_1 \longrightarrow U \longrightarrow V \longrightarrow X_2$, implies the Markov chain conditions in the outer bound in Theorem 10, i.e., (84), while (84) implies the Markov chain condition in the outer bound in Theorem 9, i.e., $X_1 \longrightarrow U \longrightarrow V$ and $U \longrightarrow V \longrightarrow X_2$.

B. New Outer Bound

We propose a new outer bound for the multiterminal rate-distortion region as follows.

Theorem 11: $\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},3}(\mathbf{D})$, where $\mathcal{R}_{\text{out},3}(\mathbf{D})$ is the set of all $\mathbf{R}$ such that there exist some positive integer n, and discrete random variables Q, X_1, X_2 for which the following three conditions are satisfied:

1) The joint distribution satisfies
$$p(u,v,x_1,x_2,q) = p(q)p(x_1,x_2|u,v,q)p(u,v) \tag{88}$$
where for given $Q = q$
$$p(x_1,x_2|u,v,Q=q) \in \mathcal{S}_{X_1X_2|UV} \tag{89}$$
with $\mathcal{S}_{X_1X_2|UV}$ defined in (37).

2) The rate pair satisfies

$$R_1 \geq I(U,V;X_1|X_2,Q) \tag{90}$$
$$R_2 \geq I(U,V;X_2|X_1,Q) \tag{91}$$
$$R_1 + R_2 \geq I(U,V;X_1,X_2|Q). \tag{92}$$

3) There exists $\big(\hat{U}(X_1,X_2,Q), \hat{V}(X_1,X_2,Q)\big)$ such that $\big(Ed_1(U,\hat{U}), Ed_2(V,\hat{V})\big) \leq \mathbf{D}$.

The proof of the this theorem can be found in Appendix III.

Next, we state and prove that our outer bound given in Theorem 11 is tighter than $\mathcal{R}_{\text{out},2}(\mathbf{D})$ given in Theorem 10.

Theorem 12:

$$\mathcal{R}_{\text{out},3}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},2}(\mathbf{D}). \tag{93}$$

Proof: We have two proofs for this theorem. We will provide the first proof here and leave the second proof to Section IV-C. We prove this theorem by construction. For every (R_1,R_2) point in $\mathcal{R}_{\text{out},3}(\mathbf{D})$, there exist random variables Q, X_1, X_2 satisfying (88), (R_1,R_2) pair satisfying (90)–(92), and a reconstruction pair $\big(\hat{U}(X_1,X_2,Q), \hat{V}(X_1,X_2,Q)\big)$ such that $\big(Ed_1(U,\hat{U}), Ed_2(V,\hat{V})\big) \leq \mathbf{D}$. According to [5], let $X_1' = (X_1,Q)$ and $X_2' = (X_2,Q)$. Then, $p(x_1',x_2'|u,v)$ satisfies (84). Moreover

$$R_1 \geq I(U,V;X_1|X_2,Q) = I(U,V;X_1'|X_2') \tag{94}$$

[7]This is a simplified version of the outer bound in [19] without introducing an extra random variable Z satisfying $Z \longrightarrow (U,V) \longrightarrow (W,X_1,X_2)$.

and similarly

$$R_2 \geq I(U,V;X_2|X_1,Q) = I(U,V;X_2'|X_1') \tag{95}$$

and finally

$$\begin{aligned} R_1 + R_2 &\geq I(U,V;X_1,X_2|Q) \\ &= H(U,V|Q) - H(U,V|X_1,X_2,Q) \\ &\overset{1}{=} H(U,V) - H(U,V|X_1,X_2,Q) \\ &= H(U,V) - H(U,V|X_1',X_2') \\ &= I(U,V;X_1',X_2') \end{aligned} \tag{96}$$

where 1 follows from the fact that Q is independent of (U,V). $(\hat{U},\hat{V})$ is a function of (X_1,X_2,Q), and, therefore, it is a function of $(X_1',X_2') = \big((X_1,Q),(X_2,Q)\big)$.

Hence, for every rate pair $(R_1,R_2) \in \mathcal{R}_{\text{out},3}(\mathbf{D})$, there exist random variables X_1', X_2' such that $p(x_1',x_2'|u_1,v_1)$ satisfies (84), (R_1,R_2) pair satisfies the mutual information constraints, and the reconstruction satisfies the distortion constraints. In other words, $(R_1,R_2) \in \mathcal{R}_{\text{out},2}(\mathbf{D})$, proving the theorem. ■

From Section II-D, we know that

$$\mathcal{S}_{X_1X_2|UV} \subseteq \mathcal{S}'_{X_1X_2|UV}. \tag{97}$$

Then, we obtain another outer bound for the multiterminal rate-distortion region as follows.

Theorem 13: $\mathcal{R}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},4}(\mathbf{D})$, where $\mathcal{R}_{\text{out},4}(\mathbf{D})$ is the set of all $\mathbf{R}$ such that there exist discrete random variable Q independent of (U,V), and discrete random variables X_1, X_2 for which the following three conditions are satisfied:

1) The joint distribution satisfies

$$p(u,v,x_1,x_2,q) = p(q)p(x_1,x_2|u,v,q)p(u,v) \tag{98}$$

where for given $Q=q$

$$p(x_1,x_2|u,v,Q=q) \in \mathcal{S}'_{X_1X_2|UV} \tag{99}$$

with $\mathcal{S}'_{X_1X_2|UV}$ defined in (44).

2) The rate pair satisfies

$$R_1 \geq I(U,V;X_1|X_2,Q) \tag{100}$$
$$R_2 \geq I(U,V;X_2|X_1,Q) \tag{101}$$
$$R_1 + R_2 \geq I(U,V;X_1,X_2|Q). \tag{102}$$

3) There exists $\big(\hat{U}(X_1,X_2,Q),\hat{V}(X_1,X_2,Q)\big)$ such that $\big(Ed_1(U,\hat{U}),Ed_2(V,\hat{V})\big) \leq \mathbf{D}$.

C. Comparison of the Bounds

All of the inner and outer bounds we discussed above are in general incomputable due to the lack of bounds on the sizes of the alphabets of the involved auxiliary random variables. Thus, we are not able to compare these bounds numerically. In this section, we will establish some relationships between these bounds by comparing the different feasible sets of the probability distributions involved in these bounds.

We begin with the inner bound. Using the time-sharing argument, a convexification of the inner bound $\mathcal{R}_{\text{in}}(\mathbf{D})$ yields another inner bound $\mathcal{R}'_{\text{in}}(\mathbf{D})$, which is larger than $\mathcal{R}_{\text{in}}(\mathbf{D})$. We define the set

$$\mathcal{S}_{\text{in}} \triangleq \{p(x_1,x_2|u,v) : X_1 \longrightarrow U \longrightarrow V \longrightarrow X_2\}. \tag{103}$$

Then, this new inner bound may be expressed as a function of $\mathcal{S}_{\text{in}}$ and $\mathbf{D}$ as follows:

$$\mathcal{R}_{\text{in}}(\mathbf{D}) \subseteq \mathcal{R}'_{\text{in}}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{in}},\mathbf{D}) \subseteq \mathcal{R}(\mathbf{D}) \tag{104}$$

where $\mathcal{F}(\mathcal{S}_{\text{in}},\mathbf{D})$ is defined as

$$\mathcal{F}(\mathcal{S}_{\text{in}},\mathbf{D}) \triangleq \bigcup_{\mathbf{p}\in\mathcal{P}(\mathcal{S}_{\text{in}},\mathbf{D})} \mathcal{C}(\mathbf{p}) \tag{105}$$

$$\begin{aligned} \mathbf{p} &\triangleq p(x_1,x_2,q|u,v) \\ &= p(x_1,x_2|u,v,Q=q)p(q) \end{aligned} \tag{106}$$

and

$$\mathcal{P}(\mathcal{S}_{\text{in}},\mathbf{D}) \triangleq \left\{ \mathbf{p} : \begin{array}{c} \forall q, p(x_1,x_2|u,v,Q=q) \in \mathcal{S}_{\text{in}}; \\ \exists\big(\hat{U}(X_1,X_2,Q),\hat{V}(X_1,X_2,Q)\big), \\ \text{s.t.}\big(Ed_1(U,\hat{U}),Ed_2(V,\hat{V})\big) \leq \mathbf{D} \end{array} \right\} \tag{107}$$

$$\mathcal{C}(\mathbf{p}) \triangleq \left\{ (R_1,R_2) : \begin{array}{r} R_1 \geq I(U,V;X_1|X_2,Q) \\ R_2 \geq I(U,V;X_2|X_1,Q) \\ R_1+R_2 \geq I(U,V;X_1,X_2|Q) \end{array} \right\}. \tag{108}$$

In [5], it was shown that $\mathcal{R}_{\text{out},1}(\mathbf{D})$ is convex. Thus, the outer bound $\mathcal{R}_{\text{out},1}(\mathbf{D})$ can be represented in terms of function $\mathcal{F}$ as well, i.e.,

$$\mathcal{R}_{\text{out},1}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},1},\mathbf{D}) \tag{109}$$

where

$$\mathcal{S}_{\text{out},1} \triangleq \left\{ p(x_1,x_2|u,v) : \begin{array}{c} X_1 \longrightarrow U \longrightarrow V \\ U \longrightarrow V \longrightarrow X_2 \end{array} \right\}. \tag{110}$$

The result by Wagner and Anatharam [19] can also be expressed by using the function $\mathcal{F}$ as

$$\mathcal{R}_{\text{out},2}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},2},\mathbf{D}) \tag{111}$$

where

$$\mathcal{S}_{\text{out},2} \triangleq \left\{ \begin{array}{l} p(x_1,x_2|u,v) : \exists w, \\ p(x_1,x_2,w|u,v) = p(w)p(x_1|w,u)p(x_2|w,v) \end{array} \right\}. \tag{112}$$

From the definition of the function $\mathcal{F}$, we can see that $\mathcal{F}$ is monotone with respect to the set argument when the distortion argument is fixed, i.e.,

$$\mathcal{F}(A,\mathbf{D}) \subseteq \mathcal{F}(B,\mathbf{D}), \quad \text{if } A \subseteq B. \tag{113}$$

Therefore, since

$$\mathcal{S}_{\text{in}} \subseteq \mathcal{S}_{\text{out},2} \subseteq \mathcal{S}_{\text{out},1} \tag{114}$$

we have

$$\begin{aligned} \mathcal{R}_{\text{in}}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{in}},\mathcal{D}) &\subseteq \mathcal{R}_{\text{out},1}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},1},\mathcal{D}) \\ &\subseteq \mathcal{R}_{\text{out},2}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},2},\mathcal{D}). \end{aligned} \tag{115}$$

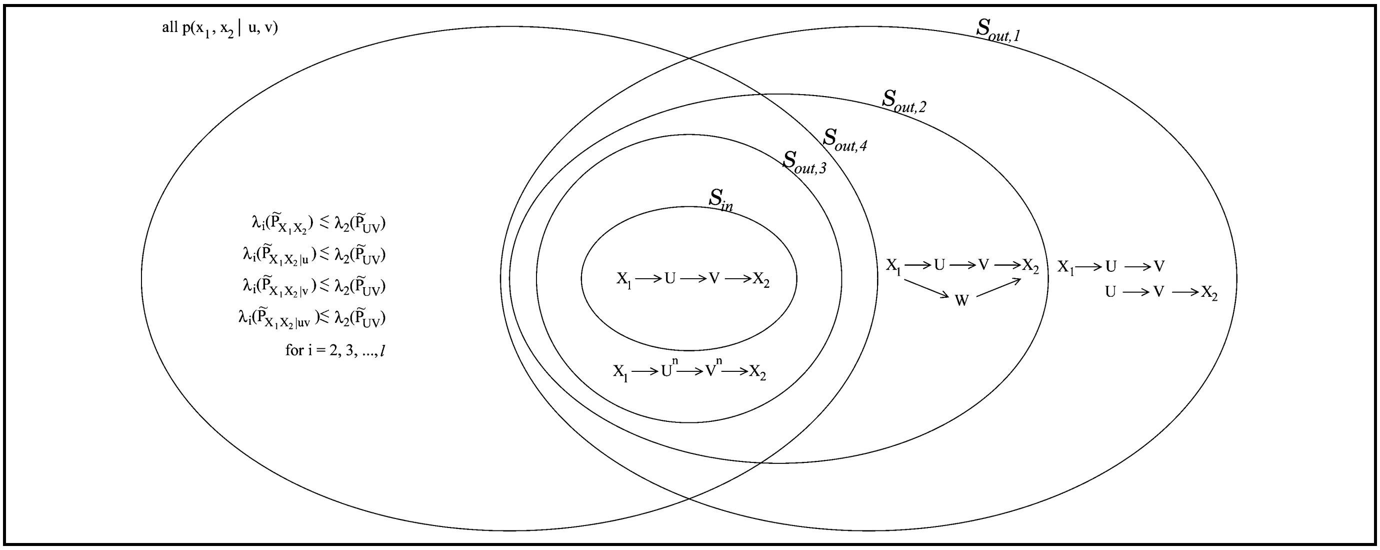

Fig. 1. Different sets of probability distributions $p(x_1, x_2|u, v)$.

We conclude that the gap between the inner and the outer bounds comes only from the difference between the feasible sets of the probability distributions $p(x_1, x_2|u, v)$.

In Theorem 12, we have shown $\mathcal{R}_{\text{out},3}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},2}(\mathbf{D})$. Here, we provide an alternative proof which comes from the comparison of the feasible sets of probability distributions $p(x_1, x_2|u, v)$. We note that $X_1 \longrightarrow U^n \longrightarrow V^n \longrightarrow X_2$ implies the Markov chain-like condition in (84) by taking $U_1 = U$, $V_1 = V$ and $(U_2^n, V_2^n) = W$, which means that

$$\mathcal{S}_{\text{out},3} \triangleq \mathcal{S}_{X_1X_2|UV} \subseteq \mathcal{S}_{\text{out},2} \tag{116}$$

and because of the monotonicity of $\mathcal{F}(\cdot, \mathbf{D})$ in (113), we have

$$\mathcal{F}(\mathcal{S}_{\text{out},3}, \mathbf{D}) = \mathcal{R}_{\text{out},3}(\mathbf{D}) \subseteq \mathcal{R}_{\text{out},2}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},2}, \mathbf{D}). \tag{117}$$

From Section II-D, we have that

$$\mathcal{S}_{\text{out},3} \subseteq \mathcal{S}_{\text{out},4} \triangleq \mathcal{S}'_{X_1X_2|UV} \tag{118}$$

and therefore

$$\mathcal{R}_{\text{out},3}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},3}, \mathbf{D}) \subseteq \mathcal{R}_{\text{out},4}(\mathbf{D}) = \mathcal{F}(\mathcal{S}_{\text{out},4}, \mathbf{D}). \tag{119}$$

So far, we have not been able to determine whether $\mathcal{S}_{\text{out},4} \subseteq \mathcal{S}_{\text{out},2}$ or $\mathcal{S}_{\text{out},2} \subseteq \mathcal{S}_{\text{out},4}$, however, we know that there exists some probability distribution $p(x_1, x_2|u_1, v_1)$, which belongs to $\mathcal{S}_{\text{out},2}$, but does not belong to $\mathcal{S}_{\text{out},4}$. For example, assume $\sigma_2(\tilde{P}_{UV}) < 1$ and some random variable W independent to (U, V). Let $X_1 = (f_1(U_1), W)$ and $X_2 = (f_2(V_1), W)$. We note that (X_1, X_2, U_1, V_1) satisfies the Markov chain-like condition in (84), i.e., $p(x_1, x_2|u_1, v_1) \in \mathcal{S}_{\text{out},2}$. But, (X_1, X_2) contains common information W, which means that $\sigma_2(\tilde{P}_{X_1X_2}) = 1 > \sigma_2(\tilde{P}_{UV})$ [10], and, therefore, $p(x_1, x_2|u_1, v_1) \notin \mathcal{S}_{\text{out},4}$. Based on this observation, we note that introducing $\mathcal{S}_{\text{out},4}$ helps us rule out some unachievable probability distributions that may exist in $\mathcal{S}_{\text{out},2}$. The relation between different feasible sets of probability distributions $p(x_1, x_2|u_1, v_1)$ is illustrated in Fig. 1.

Finally, we note that we can obtain a tighter outer bound in terms of the function $\mathcal{F}(\cdot, \mathbf{D})$ by using a set argument which is the intersection of $\mathcal{S}_{\text{out},2}$ and $\mathcal{S}_{\text{out},4}$, i.e.,

$$\mathcal{R}_{\text{out},2\cap 4}(\mathbf{D}) \triangleq \mathcal{F}(\mathcal{S}_{\text{out},2} \cap \mathcal{S}_{\text{out},4}, \mathbf{D}). \tag{120}$$

It is straightforward to see that this outer bound $\mathcal{R}_{\text{out},2\cap 4}(\mathbf{D})$ is in general tighter than the outer bound $\mathcal{F}(\mathcal{S}_{\text{out},2}, \mathbf{D})$. However, it is unknown whether our outer bound is tighter than another improvement in [19], which exploits the conditional independence in the source pair by introducing an extra random variable Z satisfying $Z \longrightarrow (U, V) \longrightarrow (X_1, X_2, W)$.

V. Conclusion

In the distributed coding on correlated sources, the problem of describing a joint distribution involving an n-letter Markov chain arises. By using a spectral method, we provided a new data processing inequality based on new measures of correlation, which gave us a single-letter necessary condition for the n-letter Markov chain. We applied our results to two specific examples involving distributed coding of correlated sources: the multi-terminal rate-distortion region and the multiple-access channel with correlated sources, and proposed two new outer bounds for these two problems.

Appendix I
Proof of Theorem 4

We consider a special case of $(\underline{U}, \underline{V})$ as follows. We define $\underline{U} \triangleq \{U_1, \ldots, U_l\}$ and $\underline{V} \triangleq \{V_1, \ldots, V_m, V_{l+1}, \ldots, V_{l+k-m}\}$. We also define the complements of $\underline{U}$ and $\underline{V}$ as: $\underline{U}^c \triangleq \{U_1, \ldots, U_n\} \backslash \underline{U}$ and $\underline{V}^c \triangleq \{V_1, \ldots, V_n\} \backslash \underline{V}$. If $\underline{U}$ and $\underline{V}$ take other forms, we can transform them to the form we defined above by permutations. We know that

$$\begin{aligned} &p(x_1, x_2, \underline{u}^c, \underline{v}^c|\underline{u}, \underline{v}) \\ &\quad = p(x_1|\underline{u}^c, \underline{u}, \underline{v}) p(\underline{u}^c, \underline{v}^c|\underline{u}, \underline{v}) p(x_2|\underline{v}^c, \underline{v}, \underline{u}). \end{aligned} \tag{121}$$

In other words, given $\underline{U} = \underline{u}$ and $\underline{V} = \underline{v}$, $(X_1, \underline{U}^c, \underline{V}^c, X_2)$ form a Markov chain. Thus, from (23), we have

$$\tilde{P}_{X_1X_2|\underline{uv}} = \tilde{P}_{X_1\underline{U}^c|\underline{uv}}\tilde{P}_{\underline{U}^c\underline{V}^c|\underline{uv}}\tilde{P}_{\underline{V}^cX_2|\underline{uv}}. \quad (122)$$

Furthermore

$$\tilde{P}_{\underline{U}^c\underline{V}^c|\underline{uv}} = \tilde{p}^T_{V^l_{m+1}|u^l_{m+1}} \otimes \tilde{p}_{U^{l+k-m}_{l+1}|v^{l+k-m}_{l+1}} \otimes \tilde{P}_{U^n_{l+k-m+1}V^n_{l+k-m+1}}. \quad (123)$$

As aforementioned, a vector marginal distribution can be viewed as a joint distribution matrix with a degenerate random variable whose alphabet size is equal to 1. Since the rank of a vector is 1, from Theorem 1, the sole singular value of $\tilde{p}_{V^l_{m+1}|u^l_{m+1}}$ (and of $\tilde{p}_{U^{l+k-m}_{l+1}|v^{l+k-m}_{l+1}}$) is equal to 1. Then

$$\sigma_i(\tilde{P}_{\underline{U}^c\underline{V}^c|\underline{uv}}) = \sigma_i(\tilde{P}_{U^n_{l+k-m+1}V^n_{l+k-m+1}}). \quad (124)$$

Combining (22), (122), and (124), we obtain

$$\sigma_i(\tilde{P}_{X_1X_2|\underline{uv}}) \le \sigma_2(\tilde{P}_{UV}) \quad (125)$$

which completes the proof.

Appendix II
Proof of Theorem 7

Consider a given block code of length n with encoders $f_1 : \mathcal{U}^n \longmapsto \mathcal{X}_1^n$ and $f_2 : \mathcal{V}^n \longmapsto \mathcal{X}_2^n$ and decoder $g : \mathcal{Y}^n \longmapsto \mathcal{U}^n \times \mathcal{V}^n$. From Fano's inequality [9, p. 39], we have

$$H(U^n, V^n|Y^n) \le n\log_2|\mathcal{U}\times\mathcal{V}|P_e + 1 \triangleq n\epsilon_n. \quad (126)$$

For a code, for which $P_e \longrightarrow 0$, as $n \longrightarrow \infty$, we have $\epsilon_n \longrightarrow 0$. Then,

$$\begin{aligned}
nH(U|V) &= H(U^n|V^n)\\
&= I(U^n;Y^n|V^n) + H(U^n|Y^n,V^n)\\
&\le I(U^n;Y^n|V^n) + H(U^n,V^n|Y^n)\\
&\overset{1}{\le} I(U^n;Y^n|V^n) + n\epsilon_n\\
&= H(Y^n|V^n) - H(Y^n|U^n,V^n) + n\epsilon_n\\
&\overset{2}{=} H(Y^n|X_2^n,V^n) - \\
&\quad - H(Y^n|X_1^n,X_2^n,U^n,V^n) + n\epsilon_n\\
&\overset{3}{=} H(Y^n|X_2^n,V^n) - H(Y^n|X_1^n,X_2^n) + n\epsilon_n\\
&\overset{4}{=} \sum_{i=1}^n \Big[H(Y_i|X_2^n,V^n,Y^{i-1}) - H(Y_i|X_{1i},X_{2i})\Big] + n\epsilon_n\\
&\overset{5}{\le} \sum_{i=1}^n \Big[H(Y_i|X_{2i},V_i) - H(Y_i|X_{1i},X_{2i})\Big] + n\epsilon_n\\
&\overset{6}{=} \sum_{i=1}^n \Big[H(Y_i|X_{2i},V_i) - H(Y_i|X_{1i},X_{2i},V_i)\Big] + n\epsilon_n\\
&= \sum_{i=1}^n I(X_{1i};Y_i|X_{2i},V_i) + n\epsilon_n
\end{aligned} \quad (127)$$

where

1) from Fano's inequality in (126);
2) from the fact that X_1^n is a deterministic function of U^n and X_2^n is a deterministic function of V^n;
3) from $p(y^n|x_1^n,x_2^n,u^n,v^n) = p(y^n|x_1^n,x_2^n)$;
4) from the chain rule and the memoryless nature of the channel;
5) from the property that conditioning reduces entropy;
6) from $p(y_i|x_{1i},x_{2i},v_i) = p(y_i|x_{1i},x_{2i})$.

Using a symmetrical argument, we obtain

$$nH(V|U) \le \sum_{i=1}^n I(X_{2i};Y_i|X_{1i},U_i) + n\epsilon_n. \quad (128)$$

Moreover,

$$\begin{aligned}
nH(U,V) &= H(U^n,V^n)\\
&= I(U^n,V^n;Y^n) + H(U^n,V^n|Y^n)\\
&\le I(U^n,V^n;Y^n) + n\epsilon_n\\
&\le I(X_1^n,X_2^n;Y^n) + n\epsilon_n\\
&= H(Y^n) - H(Y^n|X_1^n,X_2^n) + n\epsilon_n\\
&= \sum_{i=1}^n \Big[H(Y_i|Y^{i-1}) - H(Y_i|X_{1i},X_{2i})\Big] + n\epsilon_n\\
&\le \sum_{i=1}^n \Big[H(Y_i) - H(Y_i|X_{1i},X_{2i})\Big] + n\epsilon_n\\
&= \sum_{i=1}^n I(X_{1i},X_{2i};Y_i) + n\epsilon_n.
\end{aligned} \quad (129)$$

We note that the following three expressions, $I(X_{1i};Y_i|X_{2i},V_i)$, $I(X_{2i};Y_i|X_{1i},U_i)$, and $I(X_{1i},X_{2i};Y_i)$, only depend on the marginal conditional distribution $p(x_{1i},x_{2i}|u_i,v_i)$ with given $p(u_i,v_i)$ and $p(y_i|x_{1i},x_{2i})$. We also note that X_{1i} is a function of U^n and X_{2i} is a function of V^n. Thus $X_{1i} \longrightarrow U^n \longrightarrow V^n \longrightarrow X_{2i}$, and therefore $p(x_{1i},x_{2i}|u_i,v_i) \in \mathcal{S}_{X_1X_2|UV}$. Since $\mathcal{S}_{X_1X_2|UV} \subset \mathcal{S}'_{X_1X_2|UV}$, we also have $p(x_{1i},x_{2i}|u_i,v_i) \in \mathcal{S}'_{X_1X_2|UV}$.

We define a pair of random variables (U,V), which has the same distribution as the i.i.d. sources, i.e., $p(u,v) = p(u_i,v_i)$, for any $i \in \{1,\ldots,n\}$. We introduce a time-sharing random variable Q [9, p. 397] as follows. Let Q be uniformly distributed on $\{1,\ldots,n\}$ and be independent of U, V, i.e.,

$$p(u,v,q) = p(q)p(u,v). \quad (130)$$

Define random variables X_1 and X_2 to be such that

$$p(x_1,x_2|u,v,Q=i) = p(x_{1i},x_{2i}|u_i,v_i) \quad (131)$$

and $p(x_1,x_2|u,v,Q=i) \in \mathcal{S}_{X_1X_2|U_1V_1} \subset \mathcal{S}'_{X_1X_2|UV}$ for all $i = 1,\ldots,n$. Then

$$\sum_{i=1}^n I(X_{1i};Y_i|X_{2i},V_i) = nI(X_1;Y|X_2,V,Q) \quad (132)$$

$$\sum_{i=1}^n I(X_{2i};Y_i|X_{1i},U_i) = nI(X_2;Y|X_1,U,Q) \quad (133)$$

$$\sum_{i=1}^n I(X_{1i},X_{2i};Y_i) = nI(X_1,X_2;Y|Q). \quad (134)$$

Combining (132)–(134) with (127)–(129) completes the major part of proof.

It can be shown that the outer bound in Theorem 7 is equivalent to the following:

$$\mathbf{H} \in \mathcal{R}(\mathcal{S}_{X_1X_2|UV}) \triangleq \mathrm{co}\left\{ \bigcup_{\mathbf{p}\in\mathcal{S}_{X_1X_2|UV}} \mathcal{R}(\mathbf{p}) \right\} \subset \mathcal{R}(\mathcal{S}'_{X_1X_2|UV}) \quad (135)$$

where

$$\mathbf{H} \triangleq [H(U|V), H(V|U), H(U,V)] \quad (136)$$

$$\mathbf{p} \triangleq p(x_1, x_2|u,v) \quad (137)$$

$$\mathcal{R}(\mathbf{p}) \triangleq \left\{ [R_1,R_2,R_3] : \begin{array}{l} R_1 \le I(X_1;Y|X_2,V) \\ R_2 \le I(X_2;Y|X_1,U) \\ R_3 \le I(X_1,X_2;Y) \end{array} \right\} \quad (138)$$

and $\mathrm{co}\{\cdot\}$ represents the closure of the convex hull of the set argument. Thus from an argument similar to [20, Corollary 2.4, p. 278], we have $|\mathcal{Q}| \le 3$.

APPENDIX III
PROOF OF THEOREM 11

We consider an arbitrary block code of two distributed encoders and one joint decoder with reconstructions

$$(\hat{U},\hat{V})^n \triangleq \left((\hat{U},\hat{V})_1,\ldots,(\hat{U},\hat{V})_n\right) = (g_1(Y,Z),\ldots,g_n(Y,Z)) \quad (139)$$

where $Y = f_1(U^n)$ and $Z = f_2(V^n)$, such that

$$\begin{aligned} &\left(Ed_1(U^n,\hat{V}^n), Ed_2(V^n,\hat{V}^n)\right) \\ &\triangleq \left(\frac{1}{n}E\sum_{i=1}^n d_1(U_i,\hat{U}_i), \frac{1}{n}E\sum_{i=1}^n d_2(V_i,\hat{V}_i)\right) \\ &= (\Delta_1,\Delta_2) < (D_1+\epsilon, D_2+\epsilon). \end{aligned} \quad (140)$$

Here, we define $M_1 = |\mathcal{Y}|$ and $M_2 = |\mathcal{Z}|$, where $\mathcal{Y}$ and $\mathcal{Z}$ are alphabets of Y and Z, respectively.

We define the auxiliary random variables $X_{1i} = (Y, U^{i-1})$ and $X_{2i} = (Z, V^{i-1})$. Then, we have

$$\begin{aligned} \log_2 M_1 &\ge H(Y) \\ &= I(U^n,V^n;Y) \\ &\overset{1}{\ge} I(U^n,V^n;Y|Z) \\ &= \sum_{i=1}^n I(U_i,V_i;Y|Z,U^{i-1},V^{i-1}) \\ &= \sum_{i=1}^n I(U_i,V_i;Y,Z|U^{i-1},V^{i-1}) \\ &\quad - I(U_i,V_i;Z|U^{i-1},V^{i-1}) \\ &\overset{2}{=} \sum_{i=1}^n I(U_i,V_i;Y,Z|U^{i-1},V^{i-1}) \\ &\quad - I(U_i,V_i;Z|V^{i-1}) \\ &= \sum_{i=1}^n I(U_i,V_i;Y,Z,U^{i-1}|V^{i-1}) \\ &\quad - I(U_i,V_i;U^{i-1}|V^{i-1}) - I(U_i,V_i;Z|V^{i-1}) \\ &\overset{3}{=} \sum_{i=1}^n I(U_i,V_i;Y,Z,U^{i-1}|V^{i-1}) \\ &\quad - I(U_i,V_i;Z|V^{i-1}) \\ &= \sum_{i=1}^n I(U_i,V_i;Y,U^{i-1}|Z,V^{i-1}) \\ &= \sum_{i=1}^n I(U_i,V_i;X_{1i}|X_{2i}) \end{aligned} \quad (141)$$

where

1) follows from the fact that $Y \longrightarrow U^n \longrightarrow V^n \longrightarrow Z$. We observe that the equality holds when Y is independent of Z;
2) follows from the fact that
$$p(z|u_i,v_i,v^{i-1}) = p(z|u_i,v_i,u^{i-1},v^{i-1}) \quad (142)$$
3) follows from the memoryless property of the sources.

Using a symmetrical argument, we obtain

$$\log_2 M_2 \ge \sum_{i=1}^n I(U_i,V_i;X_{2i}|X_{1i}). \quad (143)$$

Moreover

$$\begin{aligned} \log_2 M_1M_2 &\ge H(Y,Z) \\ &= I(U^n,V^n;Y,Z) \\ &= \sum_{i=1}^n H(U_i,V_i) \\ &\quad - H(U_i,V_i|Y,Z,U^{i-1},V^{i-1}) \\ &= \sum_{i=1}^n I(U_i,V_i;X_{1i},X_{2i}). \end{aligned} \quad (144)$$

We define the reconstruction function as follows:

$$\begin{aligned} (\hat{U}_i,\hat{V}_i) &= g'_i(X_{1i},X_{2i}) \\ &= g'_i\left((Y,U^{i-1}),(Z,V^{i-1})\right) = g_i(Y,Z) \end{aligned} \quad (145)$$

where g_i is defined in (139). Then, the expected distortion is

$$\begin{aligned} &\left(Ed_1(U^n,\hat{V}^n), Ed_2(V^n,\hat{V}^n)\right) \\ &= \left(\frac{1}{n}\sum_{i=1}^n Ed_1(U_i,\hat{U}_i), \frac{1}{n}\sum_{i=1}^n Ed_2(V_i,\hat{V}_i)\right) \\ &= (\Delta_1,\Delta_2). \end{aligned} \quad (146)$$

We note that the three mutual information expressions, i.e., $I(U_i,V_i;X_{1i}|X_{2i})$, $I(U_i,V_i;X_{2i}|X_{1i})$, and $I(U_i,V_i;X_{1i},X_{2i})$, and the two distortion expressions, i.e., $Ed_1(U_i,\hat{U}_i)$ and $Ed_1(U_2,\hat{U}_2)$, only depend on the marginal conditional distribution $p(x_{1i},x_{2i}|u_i,v_i)$ and function g'_i with given $p(u_i,v_i)$. We also note that X_{1i} is a function of U^n and X_{2i} is a function of V^n. Thus $X_{1i} \longrightarrow U^n \longrightarrow V^n \longrightarrow X_{2i}$, and therefore $p(x_{1i},x_{2i}|u_i,v_i) \in \mathcal{S}_{X_1X_2|U_1V_1}$.

We define a pair of random variables (U,V), which has the same distribution as the i.i.d. sources, i.e., $p(u,v) = p(u_i,v_i)$, for any $i \in \{1,\ldots,n\}$. We introduce a time-sharing random variable Q, which is uniformly distributed on $\{1,\dots,n\}$ and independent of U and V, i.e.,

$$p(u,v,q)=p(u,v)p(q). \tag{147}$$

Define random variables X_1 and X_2 on the alphabet $\mathcal{X}_1\times\mathcal{X}_2=\mathcal{X}_{1j}\times\mathcal{X}_{2k}$, where $|\mathcal{X}_{1j}|=\max(|\mathcal{X}_{1i}|)$ and $|\mathcal{X}_{1k}|=\max(|\mathcal{X}_{2i}|)$ for $i=1,\dots,n$, and

$$\begin{aligned}&p(x_{1i},x_{2i}|u_i,v_i)\\ &\quad=\begin{cases}p(x_1,x_2|u,v,Q=i) & \text{for } x_1\in\mathcal{X}_{1i} \text{ and } x_2\in\mathcal{X}_{2i}\\ 0 & \text{elsewhere}\end{cases}\end{aligned} \tag{148}$$

and therefore $p(x_1,x_2|u,v,Q=i)\in\mathcal{S}_{X_1X_2|U_1V_1}$ for all $i=1,\dots,n$. Then,

$$\sum_{i=1}^{n} I(U_i,V_i;X_{1i}|X_{2i})=nI(U_1,V_1;X_1|X_2,Q) \tag{149}$$
$$\sum_{i=1}^{n} I(U_i,V_i;X_{2i}|X_{1i})=nI(U_1,V_1;X_2|X_1,Q) \tag{150}$$
$$\sum_{i=1}^{n} I(U_i,V_i;X_{1i},X_{2i})=nI(U_1,V_1;X_1,X_2|Q). \tag{151}$$

Define a reconstruction function $g(X_1,X_2,Q)=(\hat{U},\hat{V})$ to be such that

$$g(X_1,X_2,Q=i)=g_i'(X_1,X_2). \tag{152}$$

Then

$$\sum_{i=1}^{n} Ed_1(U_i,\hat{U}_i)=nEd_1(U,\hat{U})=n\Delta_1 \tag{153}$$
$$\sum_{i=1}^{n} Ed_1(V_i,\hat{V}_i)=nEd_1(V,\hat{V})=n\Delta_2. \tag{154}$$

So far we have shown that

$$\begin{aligned}&(R_1+\delta,R_2+\delta)\\ &\quad=\left(\frac{1}{n}\log_2 M_1,\frac{1}{n}\log_2 M_2\right)\in R_{\text{out},3}\left((\Delta_1,\Delta_2)\right).\end{aligned} \tag{155}$$

We know that $(\Delta_1,\Delta_2)\le(D_1+\epsilon,D_2+\epsilon)$. Because of the monotonicity of the function $R_{\text{out},3}(\cdot)$, we have

$$\begin{aligned}(R_1+\delta,R_2+\delta)&=\left(\frac{1}{n}\log_2 M_1,\frac{1}{n}\log_2 M_2\right)\\ &\in R_{\text{out},3}\big((\Delta_1,\Delta_2)\big)\\ &\subseteq R_{\text{out},3}\big((D_1+\epsilon,D_2+\epsilon)\big).\end{aligned} \tag{156}$$

Let $\delta\longrightarrow 0$ and $\epsilon\longrightarrow 0$. Due to the continuity of the function $R_{\text{out},3}(\cdot)$, which will be proven in Appendix IV, we have [4], [18]

$$(R_1,R_2)\in R_{\text{out},3}\big((D_1,D_2)\big). \tag{157}$$

Appendix IV
Some Properties of Function $\mathcal{F}$

Function $\mathcal{F}(\cdot,\cdot)$ has two arguments, the probability set argument and the distortion argument. We recall the definition of $\mathcal{F}$ as follows:

$$\mathcal{F}(\mathcal{S},\mathbf{D})\triangleq\bigcup_{\mathbf{p}\in\mathcal{P}(\mathcal{S},\mathbf{D})}\mathcal{C}(\mathbf{p}) \tag{158}$$
$$\mathbf{p}\triangleq p(x_1,x_2,q|u,v)=p(x_1,x_2|u,v,Q=q)p(q) \tag{159}$$
$$\mathcal{P}(\mathcal{S},\mathbf{D})\triangleq\left\{\mathbf{p}:\begin{array}{c}\forall q, p(x_1,x_2|u,v,Q=q)\in\mathcal{S}_{\text{in}};\\ \exists\big(\hat{U}(X_1,X_2,Q),\hat{V}(X_1,X_2,Q)\big),\\ \text{s.t. }\big(Ed_1(U,\hat{U}),Ed_2(V,\hat{V})\big)\le\mathbf{D}\end{array}\right\} \tag{160}$$
$$\mathcal{C}(\mathbf{p})\triangleq\left\{(R_1,R_2):\begin{array}{r}R_1\ge I(U,V;X_1|X_2,Q)\\ R_2\ge I(U,V;X_2|X_1,Q)\\ R_1+R_2\ge I(U,V;X_1,X_2|Q)\end{array}\right\}. \tag{161}$$

From the definition, we note that for the probability set argument, if $A\subseteq B$, then

$$\mathcal{P}(A,\mathbf{D})\subseteq\mathcal{P}(B,\mathbf{D}) \tag{162}$$

and therefore

$$\mathcal{F}(A,\mathbf{D})=\bigcup_{\mathbf{p}\in\mathcal{P}(A,\mathbf{D})}\mathcal{C}(\mathbf{p})\subseteq\mathcal{F}(B,\mathbf{D})=\bigcup_{\mathbf{p}\in\mathcal{P}(B,\mathbf{D})}\mathcal{C}(\mathbf{p}) \tag{163}$$

which means that function $\mathcal{F}$ is monotone in the probability set argument.

Similarly, if $\mathbf{D}_1\le\mathbf{D}_2$, then

$$\mathcal{P}(\mathbf{S},\mathbf{D}_1)\subseteq\mathcal{P}(\mathbf{S},\mathbf{D}_2) \tag{164}$$

and therefore

$$\mathcal{F}(\mathbf{S},\mathbf{D}_1)=\bigcup_{\mathbf{p}\in\mathcal{P}(\mathbf{S},\mathbf{D}_1)}\mathcal{C}(\mathbf{p})\subseteq\mathcal{F}(\mathbf{S},\mathbf{D}_2)=\bigcup_{\mathbf{p}\in\mathcal{P}(\mathbf{S},\mathbf{D}_2)}\mathcal{C}(\mathbf{p}) \tag{165}$$

which means that function $\mathcal{F}$ is monotone in the distortion argument.

Consider two distortions $\mathbf{D}_1$ and $\mathbf{D}_2$ such that

$$\mathbf{D}=\lambda\mathbf{D}_1+(1-\lambda)\mathbf{D}_2 \tag{166}$$

where $0\le\lambda\le 1$. Assume rate pairs $\mathbf{R}_1\in\mathcal{F}(\mathcal{S},\mathbf{D}_1)$ and $\mathbf{R}_2\in\mathcal{F}(\mathcal{S},\mathbf{D}_2)$. We note that there exists $\mathbf{p}_1\in\mathcal{P}(\mathbf{S},\mathbf{D}_1)$ and $\mathbf{p}_2\in\mathcal{P}(\mathbf{S},\mathbf{D}_2)$ such that $\mathbf{R}_1\in\mathcal{C}(\mathbf{p}_1)$ and $\mathbf{R}_2\in\mathcal{C}(\mathbf{p}_2)$. We define a binary random variable Λ with $Pr(\Lambda=1)=\lambda$ and $Pr(\Lambda=2)=1-\lambda$ and we define $Q'=(Q,\Lambda)$ and $\mathbf{p}\triangleq p(x_1,x_2,q'|u,v)$, where

$$p(x_1,x_2,q,\Lambda=i|u,v)=p_i(x_1,x_2,q|u,v),\qquad i=1,2. \tag{167}$$

It is easy to check that $\mathbf{R} \triangleq \lambda\mathbf{R}_1 + (1-\lambda)\mathbf{R}_2 \in \mathcal{C}(\mathbf{p})$ and $p \in \mathcal{P}(\mathbf{S},\mathbf{D})$. Thus

$$\mathbf{R} \in \mathcal{F}(\mathcal{S},\mathbf{D}) \tag{168}$$

i.e., $\mathcal{F}$ is convex in the distortion argument.

By a similar argument, we can show that if $\mathbf{R}_1$ and $\mathbf{R}_2$ are both in the set $\mathcal{F}(\mathcal{S},\mathbf{D})$, then $\mathbf{R} \triangleq \lambda\mathbf{R}_1 + (1-\lambda)\mathbf{R}_2 \in \mathcal{F}(\mathcal{S},\mathbf{D})$, i.e., $\mathcal{F}(\mathcal{S},\mathbf{D})$ is a convex set.

Finally, we will show the continuity of $\mathcal{F}(\mathcal{S},\mathbf{D})$. We assume that $\mathcal{S}$ includes the conditional probability corresponding to the deterministic case where $X_1 = U$ and $X_2 = V$. In this case, $\mathcal{F}(\mathcal{S},\mathbf{0})$ is inner bounded by the Slepian-Wolf region. Due to the the monotonicity of $\mathcal{F}(\mathcal{S},\mathbf{D})$ in $\mathbf{D}$, the boundary of $\mathcal{F}(\mathcal{S},\mathbf{D})$ for any $\mathbf{D}$ lies outside of the Slepian-Wolf region. We also note that for any point on the boundary of $\mathcal{F}(\mathcal{S},\mathbf{D})$, the distance between this point and the Slepian-Wolf region is upper bounded by a finite number, say l, where the distance here is the Euclidean distance in two-dimansional space, and therefore, the distance between this point and $\mathcal{F}(\mathcal{S},(D_1-a,D_2-a))$ with $0 < a < \min(D_1,D_2)$ is also upper bounded by l. Because of the convexity of $\mathcal{F}(\mathcal{S},\mathbf{D})$ in $\mathbf{D}$, the distance between this point and $\mathcal{F}(\mathcal{S},(D_1-\epsilon,D_2-\epsilon))$ with $\epsilon < \alpha$ is upper bounded by $\frac{\epsilon l}{\alpha}$, which proves the continuity of $\mathcal{F}(\mathcal{S},\mathbf{D})$.

References

[1] W. Kang and S. Ulukus, "A single-letter upper bound for the sum rate of multiple access channels with correlated sources," in *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2005.

[2] W. Kang and S. Ulukus, "An outer bound for multiple access channels with correlated sources," in *Proc. 40th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2006.

[3] W. Kang and S. Ulukus, "An outer bound for multiterminal rate-distortion region," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, WA, Jul. 2006.

[4] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, Cornell Univ., Ithaca, NY, 1978.

[5] K. B. Housewright, "Source coding studies for multiterminal systems," Ph.D. dissertation, Univ. California, Los Angeles, CA, 1977.

[6] T. Berger, *Multi-Terminal Source Coding. Chapter in The Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1978.

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

[8] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 396–412, May 1983.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[10] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, pp. 100–113, Jan. 1975.

[11] K. Marton, "The structure of isomorphisms of discrete memoryless correlated sources," *Zeitschrift Wahrscheinlichkeitstheorie Verwandte Gebiete*, vol. 56, no. 3, pp. 317–327, Jan. 1981.

[12] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic, 1979.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[15] W. Kang, "Correlation and cooperation in network information theory," Ph.D. dissertation, Univ. Maryland, College Park, 2008.

[16] G. Dueck, "A note on the multiple access channel with correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 2, pp. 232–235, Mar. 1981.

[17] S. S. Pradhan, S. Choi, and K. Ramchandran, "A graph-based framework for transmission of correlated sources over multiple access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4583–4604, Dec. 2007.

[18] 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.

[19] A. B. Wagner and V. Anantharam, "An improved outer bound for the multiterminal source coding problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.

[20] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

Wei Kang (S'05–M'08) received the B.Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2001, the M.Eng. degree in electrical engineering from McGill University, Montreal, Canada, in 2003, and the Ph.D. degree from the University of Maryland, College Park, in 2008.

He joined the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2009, where he is currently an Associate Research Professor. His research interests include network information theory.

Sennur Ulukus (S'90–M'98) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1991 and 1993, respectively, and the Ph.D. degree in electrical and computer engineering from Rutgers University, New Brunswick, NJ, in 1998.

During her Ph.D. studies, she was with the Wireless Information Network Laboratory (WINLAB), Rutgers University. From 1998 to 2001, she was a Senior Technical Staff Member with AT&T Labs-Research, NJ. In 2001, she joined the University of Maryland at College Park, where she is currently an Associate Professor with the Department of Electrical and Computer Engineering, with a joint appointment at the Institute for Systems Research (ISR). Her research interests are in wireless communication theory and networking, network information theory for wireless networks, signal processing for wireless communications, and security for multiuser wireless communications.

Dr. Ulukus is a recipient of the 2005 NSF CAREER Award, and a corecipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications. She serves/served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY since 2007, as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 2003–2007, as a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY for the Special Issue on Interference Networks, as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the Special Issue on Multiuser Detection for Advanced Communication Systems and Networks. She has been the Co-Chair of the following conferences: Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference; Medium Access Control (MAC) Track at the 2008 IEEE Wireless Communications and Networking Conference; Wireless Communications Symposium at the 2010 IEEE International Conference on Communications; 2011 Communication Theory Workshop; Physical-Layer Security Workshop at the 2011 IEEE International Conference on Communications; and was the Secretary of the IEEE Communication Theory Technical Committee (CTTC) in 2007–2009.