arXiv:0807.2666v1 [cs.IT] 16 Jul 2008

Transmission of Correlated Sources Over Multi-user Channels

Deniz Gündüz, Elza Erkip, Andrea Goldsmith, H. Vincent Poor

Abstract

Source and channel coding over multi-user channels in which receivers have access to correlated source side information, is considered. For several simple multi-user channels necessary and sufficient conditions are characterized for lossless transmission under certain source-channel structures. In particular, the multiple access channels, the compound multiple access channels and the interference channels with correlated sources and correlated receiver side information are considered. In all of the cases for which the optimal source-channel matching conditions have been identified, this optimal performance is shown to be achieved by separately designed source and channel codes.

Index Terms

Network information theory, source-channel separation theorem, source and channel coding.

The material in this paper was presented in part at the 2nd UCSD Information Theory and Applications Workshop (ITA), San Diego, CA, Jan. 2007, at the Data Compression Conference (DCC), Snowbird, UT, March 2007, and at the IEEE International Symposium on Information Theory (ISIT), Nice, France, June 2007.

This research was supported in part by the U.S. National Science Foundation under Grants ANI-03-38807, CCF-04-30885, CCF-06-35177, CCF-07-28208, and CNS-06-25637 and DARPA ITMANET program under grant 1105741-1-TFIND and the ARO under MURI award W911NF-05-1-0246.

Deniz Gündüz is with the Department of Electrical Engineering, Princeton University, Princeton, NJ, 08544 and with the Department of Electrical Engineering, Stanford University, Stanford, CA, 94305 (email: dgunduz@princeton.edu).

Elza Erkip is with the Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY 11201 (email: elza@poly.edu).

Andrea Goldsmith is with the Department of Electrical Engineering, Stanford University, Stanford, CA, 94305 (email: andrea@systems.stanford.edu).

H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ, 08544 (email: poor@princeton.edu).

I. Introduction

Shannon's source-channel separation theorem states that, in point-to-point communication systems (under certain conditions), a source can be reliably transmitted over a channel if and only if the minimum source coding rate is below the channel capacity, and that the source and channel codes can be designed independently. This means that a simple comparison of the rates of the optimal source and channel codes suffices to conclude whether reliable transmission is possible or not. Furthermore, modularity in communication system design does not incur any loss in terms of the end-to-end system performance. This theoretical optimality of modularity has led to the layered design for point-to-point systems, and to the separate development of source and channel coding research. The separation theorem holds for ergodic sources and channels under the usual information theoretic assumptions of infinite delay and complexity (see [2] for more general conditions under which separation holds). However; Shannon's source-channel separation theorem does not generalize to multi-user networks.

Suboptimality of separation for multi-user systems was first shown by Shannon in [3] where he provided an example of correlated source transmission over the two-way channel. Later, a similar observation was made for transmitting correlated sources over multiple access channels (MACs) in [4]. The example provided in [4] reveals that comparison of the Slepian-Wolf source coding region [5] with the capacity region of the underlying MAC is not enough to decide whether reliable transmission can be realized.

In the most general setting, a source-channel network problem is composed of multiple sources available at the nodes of a network in which the source data must be transmitted to its destination in a lossless or lossy fashion. Some (potentially all) of the nodes can transmit while some (potentially all) of the nodes can receive noisy observations of the transmitted signals. The communication channel is characterized by a probability transition matrix from the inputs of the transmitting terminals to the outputs of the receiving terminals. We assume that all the transmissions share a common communications medium, but other special cases can be specified through the channel transition matrix. The sources come from an arbitrary joint distribution, that is, they might be correlated. For this very broad model, the general problem we address is to determine whether the sources can be transmitted losslessly or within the required fidelity to their destinations for a given number of channel uses per source sample, which is defined to be the *source-channel rate* of the joint source channel code. Equivalently, we might want to find the minimum source-channel rate that can achieve either reliable transmission (for lossless reconstruction) or the required reconstruction fidelity (for lossy reconstruction).

This general setup can model a sensor network application where multiple sensors spread over a region of interest send their correlated observations to an access point over a multiple access channel. However, the problem in this very general setting is extremely complicated. If the channels are assumed to be noise-free finite capacity links, the problem reduces to a multiterminal source coding problem [6]; alternatively, if the sources are independent, then we are faced with finding the capacity of a general communication network. Furthermore, considering that we do not have a separation result even in the case of very simple networks, the hope to resolve the above source-channel problem in the general setting is slight.

Given the difficulty of obtaining a general solution for arbitrary networks, the main goal in joint source-channel coding for multi-user systems is to analyze in detail simple, yet fundamental, building blocks of a larger network, such as the multiple access channel, the broadcast channel and the interference channel. Our focus in this work is on lossless transmission and our goal is to characterize the set of achievable source-channel rates for these canonical networks. Three fundamental questions that need to be addressed for each model can be stated as follows:

1) Is it possible to characterize the optimal source-channel rate of the network (i.e., the minimum number of channel uses per source sample required for lossless transmission) in a computable way?
2) For a given channel model, what is the most general class of sources such that the optimal source-channel rate can be characterized by simply comparing the achievable source coding rate region with the achievable capacity region?
3) When the comparison of these regions is not sufficient to obtain the optimal source-channel rate, what is the maximum source-channel rate achievable by source and channel codes that are statistically independent? How does this rate compare to the optimal?

If the class of the sources in question (2) corresponds to all the sources of interest for a given setup, this would maintain the optimality of the layered approach described earlier. However, even when this layered approach is suboptimal, we can still obtain modularity in the system design, if the optimal rate can be achieved by statistically independent source and channel codes as in question (3). Statistical independence in this setup refers to designing source and channel codes solely depending on the source and channel distributions, respectively, without taking the joint distribution into account.

The existing literature provides limited answers to the above three questions in specific settings. For the MAC with correlated sources, finite-letter sufficient conditions for achievability of a source-channel rate are given in [4] in an attempt to resolve the first problem; however, these conditions are later shown not to be necessary by Dueck [7]. The *correlation preserving mapping* technique of [4] used for achievability is later extended to source coding with side information via multiple access channels in [8], to broadcast channels with correlated sources in [9], and to interference channels in [10]. In [11], [12] a graph theoretic framework was used to achieve improved rates for transmitting correlated sources over multiple access and broadcast channels, respectively. A new data processing inequality was proven in [13] that is used to derive new necessary conditions for reliable transmission of correlated sources over MACs.

In an effort to resolve the second question above, which asks for the most general class of sources for which the rate region comparison is sufficient to determine the achievability of a source-channel rate, various special classes of source-channel pairs have been studied in the literature. Optimality of separation for a network of independent, non-interfering channels is proven in [14]. A special class of the MAC, called the asymmetric MAC, in which one of the sources is available at both encoders, is considered in [15], and source-channel separation is shown to hold with or without causal perfect feedback at either or both of the transmitters. In [16], it is shown that for the class of MACs for which the capacity region cannot be enlarged by considering correlated channel inputs, separation is optimal. Note that in all of these results, a special class of MACs is considered while arbitrary source correlation is allowed.

In [17], Tuncel finds the optimum rate for broadcasting a common source to multiple receivers having different correlated side information. Apart from answering the first question, which asks for the minimum source-channel rate, for this model of interest, [17] also answers the third question, and suggests that we can have optimal codes that are statistically independent of each other even though the comparison of the broadcast channel capacity region and the minimum source coding rate region is not sufficient to decide whether reliable transmission is possible. The codes proposed in [17] consist of a source encoder that does not use the correlated side information, and a joint source-channel decoder; hence they cannot operate as stand-alone source and channel codes. Doing so would require the design of new coding schemes; however, it is shown in [18] that the same performance can be achieved by using separate source and channel codes with a specific message passing mechanism between the source/channel encoders/decoders. Hence we can use existing near-optimal codes to achieve the theoretical bound.

Following [17] one can define a generalization of Shannon's source and channel separation theorem. *Informational separation* refers to separation in the Shannon sense, in which concatenating optimal source and channel codes for the underlying source and channel structures result in the optimal source-channel coding rate. *Operational separation*, on the other hand, requires only that the source and the channel codes be statistically independent, while individually they may not be suitable for source or channel coding. Note that, the class of codes satisfying operational separation is larger than that satisfying informational separation.

Our goal in this paper is to provide answers to the three fundamental questions about source-channel coding posed above for some special multi-user networks and source structures. In particular, we consider correlated sources available at multiple users to be transmitted to receivers with correlated side information. Our contributions can be summarized as follows.

- In a multiple access channel we show that informational separation holds if the sources are independent given the receiver side information. This is different from the previous separation results [16], [14] in that we show the optimality of separation for an arbitrary multiple access channel under a special source structure.
- We characterize an achievable source-channel rate for compound multiple access channels, which is shown to be optimal for some special scenarios. In particular, optimality holds either when each user's source is correlated with one of the receiver side information and is independent from the other, or when there is no multiple access interference at the receivers. We further show the optimality of informational separation when the two sources are independent given the side information common to both receivers. Note that the compound multiple access channel model combines both the multiple access channels with correlated sources and the broadcast channels with correlated side information at the receivers.
- For the interference channels, we first define the *strong source-channel interference* conditions, which provide a generalization of the usual strong interference conditions [9]. Our results suggest the optimality of operational separation under strong source-channel interference conditions for certain source structures.

Overall, our results characterize the necessary and sufficient conditions for reliable transmission of correlated sources for various multi-user scenarios, hence answering question (1) for those scenarios. In all these cases, the best performance is achieved by source-channel separation (either informational or operational), thus promising a level of modularity even when simply concatenating optimal source and channel codes is suboptimal. Hence, for the cases considered, we answer questions (2) and (3) as well by showing that, either comparing source and channel coding rate regions is sufficient to characterize the optimal source-channel rate, i.e., informational separation, or simple region comparison is not sufficient but we can build separate source and channel codes that are optimal, i.e., operational separation.

II. Preliminaries

A. Notation

In the rest of the paper we adopt the following notational conventions. Random variables will be denoted by capital letters while their realizations will be denoted by the respective lower case letters. The alphabet of a scalar random variable X will be denoted by the corresponding calligraphic letter $\mathcal{X}$, and the alphabet of the n-length vectors over the n-fold Cartesian product by $\mathcal{X}^n$. The cardinality of set $\mathcal{X}$ will be denoted by $|\mathcal{X}|$. The random vector $(X_1, \ldots, X_n)$ will be denoted by X^n while the vector $(X_i, X_{i+1}, \ldots, X_n)$ by X_i^n, and their realizations, respectively, by $(x_1, \ldots, x_n)$ or x^n and $(x_i, x_{i+1}, \ldots, x_n)$ or x_i^n.

B. Types and Typical Sequences

Here, we briefly review the notions of types and strong typicality that will be used in the paper. Given a distribution p_X, the type P_{x^n} of an n-tuple x^n is the empirical distribution

$$P_{x^n} = \frac{1}{n} N(a|x^n)$$

where $N(a|x^n)$ is the number of occurances of the letter a in x^n. The set of all n-tuples x^n with type Q is called the type class Q and denoted by T_Q^n. The set of δ-strongly typical n-tuples according to P_X is denoted by $T_{[X]_\delta}^n$ and is defined by

$$T_{[X]_\delta}^n = \left\{ x \in \mathcal{X}^n : \left| \frac{1}{n} N(a|x^n) - P_X(a) \right| \leq \delta \forall a \in \mathcal{X} \text{ and } N(a|x^n) = 0 \text{ whenever } P_X(x) = 0 \right\}.$$

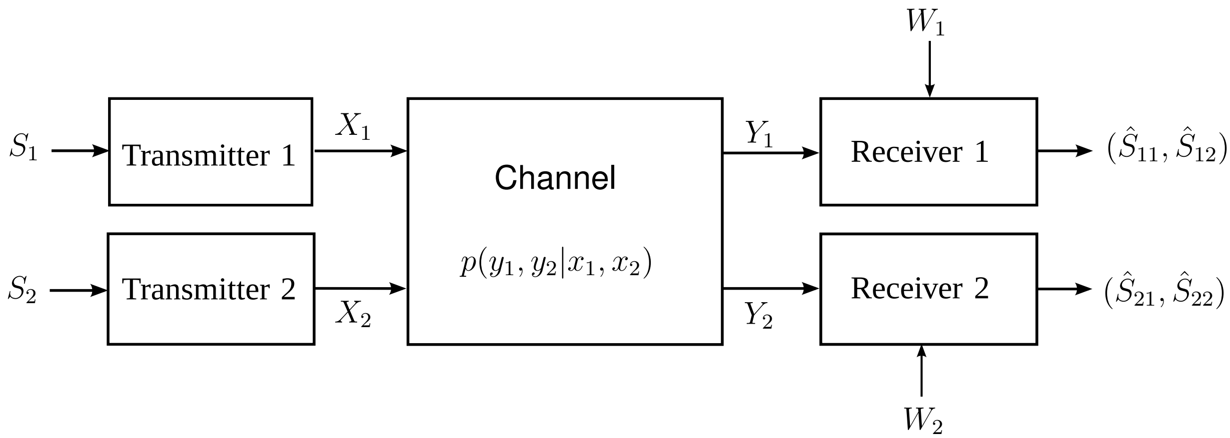

Fig. 1. The general model considered in the paper for transmitting correlated sources over multi-user channels to receivers with correlated side information. In the MAC scenario, we have only one receiver Rx_1, which wants to receive both sources S_1 and S_2. In the compound MAC scenario, we have two receivers which want to receive both sources, while in the interference channel scenario, we have two receivers, each of which wants to receive only its own transmitter's source.

The definitions of type and strong typicality can be extended to joint and conditional distributions in a similar manner [6]. The following results concerning typical sets will be used in the sequel. We have

$$\left| \frac{1}{n} \log |T^n_{[X]_\delta}| - H(X) \right| \leq \delta \tag{1}$$

for sufficiently large n. Given a joint distribution P_{XY}, if $(x^n, y^n) \sim P^n_X P^n_Y$, where P^n_X and P^n_Y are n-fold products of the marginals P_X and P_Y, then

$$\Pr\{(x^n, y^n) \in T^n_{[XY]_\delta}\} \leq 2^{-n(I(X;Y)-3\delta)}. \tag{2}$$

Finally, for a joint distribution P_{XYZ}, if $(x^n, y^n, z^n) \sim P^n_X P^n_Y P^n_Z$ with similar definitions for P^n_X, P^n_Y and P^n_Z, then

$$\Pr\{(x^n, y^n, z^n) \in T^n_{[XYZ]_\delta}\} \leq 2^{-n(I(X;Y,Z)+I(Y;X,Z)+I(Z;Y,X)-4\delta)}. \tag{3}$$

III. System Model

We introduce the most general system model here. Throughout the paper we consider various special cases, where the restrictions are stated explicitly for each case.

We consider a network of two transmitters Tx_1 and Tx_2, and two receivers Rx_1 and Rx_2. For $i = 1, 2$, the transmitter Tx_i observes the output of a discrete memoryless (DM) source S_i, while the receiver Rx_i observes DM side information W_i. We assume that the source and the side information sequences, $\{S_{1,k}, S_{2,k}, W_{1,k}, W_{2,k}\}_{k=1}^{\infty}$ are independent identically distributed (i.i.d.) and are drawn according to a joint probability mass function (p.m.f.) $p(s_1, s_2, w_1, w_2)$ over a finite alphabet $\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{W}_1 \times \mathcal{W}_2$. The transmitters and the receivers all know this joint p.m.f., but have no direct access to each other's information source or the side information.

The transmitter Tx_i encodes its source vector $S_i^m = (S_{i,1}, \dots, S_{i,m})$ into a channel codeword $X_i^n = (X_{i,1}, \dots, X_{i,n})$ using the encoding function

$$f_i^{(m,n)} : \mathcal{S}_i^m \to \mathcal{X}_i^n, \tag{4}$$

for $i = 1, 2$. These codewords are transmitted over a DM channel to the receivers, each of which observes the output vector $Y_i^n = (Y_{i,1}, \dots, Y_{i,n})$. The input and output alphabets $\mathcal{X}_i$ and $\mathcal{Y}_i$ are all finite. The DM channel is characterized by the conditional distribution $P_{Y_1,Y_2|X_1,X_2}(y_1, y_2|x_1, x_2)$.

Each receiver is interested in one or both of the sources depending on the scenario. Let receiver Rx_i form the estimates of the source vectors S_1^m and S_2^m, denoted by $\hat{S}_{i,1}^m$ and $\hat{S}_{i,2}^m$, based on its received signal Y_i^n and the side information vector $W_i^m = (W_{i,1}, \dots, W_{i,m})$ using the decoding function

$$g_i^{(m,n)} : \mathcal{Y}_i^n \times \mathcal{W}_i^m \to \mathcal{S}_1^m \times \mathcal{S}_2^m. \tag{5}$$

Due to the reliable transmission requirement, the reconstruction alphabets are the same as the source alphabets. In the MAC scenario, there is only one receiver Rx_1, which wants to receive both of the sources S_1 and S_2. In the compound MAC scenario, both receivers want to receive both sources, while in the interference channel scenario, each receiver wants to receive only its own transmitter's source. Based on these decoding requirements, the error probability of the system will be defined separately for each model. Next, we define the source-channel rate of the system.

Definition 3.1: We say that source-channel rate b is *achievable* if, for every $\epsilon > 0$, there exist positive integers m and n with $n/m = b$ for which we have encoders $f_1^{(m,n)}$ and $f_2^{(m,n)}$, and decoders $g_1^{(m,n)}$ and $g_2^{(m,n)}$ with decoder outputs $(\hat{S}_{i,1}^m, \hat{S}_{i,2}^m) = g_i(Y_i^n, W_i^m)$, such that $P_e^{(m,n)} < \epsilon$.

IV. Multiple Access Channel

We first consider the multiple access channel, in which we are interested in the reconstruction at receiver Rx_1 only. For encoders $f_i^{(m,n)}$ and a decoder $g_1^{(m,n)}$, the probability of error for the MAC is defined as follows:

$$\begin{aligned} P_e^{(m,n)} &\triangleq Pr\{(S_1^m, S_2^m) \neq (\hat{S}_{1,1}^m, \hat{S}_{1,2}^m)\} \\ &= \sum_{(s_1^m, s_2^m) \in \mathcal{S}_1^m \times \mathcal{S}_2^m} p(s_1^m, s_2^m) P\{(\hat{s}_{1,1}^m, \hat{s}_{1,2}^m) \neq (s_1^m, s_2^m) | (S_1^m, S_2^m) = (s_1^m, s_2^m)\}. \end{aligned}$$

Note that this model is more general than that considered in [4] as it considers the availability of correlated side information at the receiver. We first generalize the achievability scheme of [4] to our model by using the correlation preserving mapping technique of [4], and limiting the rate b to 1. Extension to other rates is possible as in Theorem 4 of [4].

Theorem 4.1: Consider arbitrarily correlated sources S_1 and S_2 over the DM MAC with receiver side information W_1. Rate $b = 1$ is achievable if

$$\begin{aligned} H(S_1|S_2, W_1) &< I(X_1; Y_1|X_2, S_2, W_1, Q), \\ H(S_2|S_1, W_1) &< I(X_2; Y_1|X_1, S_1, W_1, Q), \\ H(S_1, S_2|W, W_1) &< I(X_1, X_2; Y_1|U, W_1, Q), \end{aligned}$$

and

$$H(S_1, S_2|W_1) < I(X_1, X_2; Y_1|W_1),$$

for some joint distribution

$$p(q, s_1, s_2, w_1, x_1, x_2, y_1) = p(q)p(s_1, s_2, w_1)p(x_1|q, s_1)p(x_2|q, s_2)p(y_1|x_1, x_2)$$

and

$$U = f(S_1) = g(S_2)$$

is the common part of S_1 and S_2 in the sense of Gàcs and Körner [21]. We can bound the cardinality of Q by $\min\{|\mathcal{X}_1| \cdot |\mathcal{X}_2|, |\mathcal{Y}|\}$.

We do not give a proof here as it closely resembles the one in [4]. Note that correlation among the sources and the side information both condenses the left hand sides of the above inequalities, and enlarges their right hand sides, compared to transmitting independent sources. While the reduction in entropies on the left hand sides is due to Slepian-Wolf source coding, the increase in the right hand sides is mainly due to the possibility of generating correlated channel codewords at the transmitters. Applying distributed source coding followed by MAC channel coding, while reducing the redundancy, would also lead to the loss of possible correlation among the channel codewords. However, when $S_1 - W_1 - S_2$ form a Markov chain, that is, the two sources are independent given the side information at the receiver, the receiver already has access to the correlated part of the sources and it is not clear whether additional channel correlation would help. The following theorem suggests that channel correlation preservation is not necessary in this case and source-channel separation in the informational sense is optimal.

Theorem 4.2: Consider transmission of arbitrarily correlated sources S_1 and S_2 over the DM MAC with receiver side information W_1, for which the Markov relation $S_1 - W_1 - S_2$ holds. Rate b is achievable if

$$H(S_1|W_1) < b \cdot I(X_1;Y_1|X_2,Q), \tag{6}$$

$$H(S_2|W_1) < b \cdot I(X_2;Y_1|X_1,Q), \tag{7}$$

and

$$H(S_1|W_1) + H(S_2|W_1) < b \cdot I(X_1,X_2;Y_1|Q), \tag{8}$$

for some joint distribution

$$p(q,x_1,x_2,y_1) = p(q)p(x_1|q)p(x_2|q)p(y_1|x_1,x_2), \tag{9}$$

with $|\mathcal{Q}| \leq 4$.

Conversely, if rate b is achievable, then (6)-(8) hold with $<$ replaced by $\leq$ for some joint distribution of the form given in (9).

Proof: We start with the proof of the direct part. Consider a rate pair (R_1, R_2) satisfying

$$H(S_1|W_1) < R_1 < b \cdot I(X_1;Y_1|X_2,Q), \tag{10}$$

$$H(S_2|W_1) < R_2 < b \cdot I(X_2;Y_1|X_1,Q), \tag{11}$$

and

$$H(S_1|W_1) + H(S_2|W_1) < R_1 + R_2 < b \cdot I(X_1,X_2;Y_1|Q). \tag{12}$$

While the left hand sides of (10) and (11) form sufficient conditions for lossless compression of the sources S_1 and S_2 at rates R_1 and R_2, respectively, given a receiver with side information W_1, i.e., (R_1, R_2) is in the Slepian-Wolf rate region [5], the right hand sides of (10)-(12) constitute sufficient conditions for reliable transmission of the compressed messages over the multiple access channel, guaranteeing the achievability of rate b.

We next prove the converse. We assume $P_e^{(m,n)} \to 0$ for a sequence of encoders $f_i^{(m,n)}$ $(i = 1, 2)$ and decoders $g^{(m,n)}$ as $n, m \to \infty$ with a fixed rate $b = n/m$. We will use Fano's inequality, which states

$$\begin{aligned} H(S_1^m, S_2^m | \hat{S}_1^m, \hat{S}_2^m) &\leq 1 + mP_e^{(m,n)} \log |\mathcal{S}_1 \times \mathcal{S}_2|, \\ &\triangleq m\delta(P_e^{(m,n)}), \end{aligned} \tag{13}$$

where $\delta(P_e^{(m,n)})$ is a non-negative function that approaches to zero as $P_e^{(m,n)} \to 0$. We also obtain

$$H(S_1^m, S_2^m | \hat{S}_1^m, \hat{S}_2^m) \geq H(S_1^m | \hat{S}_1^m, \hat{S}_2^m), \tag{14}$$

$$\geq H(S_1^m | Y_1^n, W_1^m), \tag{15}$$

where the first inequality follows from the chain rule of entropy and the nonnegativity of the entropy function for discrete sources, and the second inequality follows from the data processing inequality. Then we have, for $i = 1, 2$,

$$H(S_i^m | Y_1^n, W_1^m) \leq m\delta(P_e^{(m,n)}). \tag{16}$$

We have

$$\frac{1}{n} I(X_1^n; Y_1^n | X_2^n, W_1^m) \geq \frac{1}{n} I(S_1^m; Y_1^n | W_1^m, X_2^n), \tag{17}$$

$$= \frac{1}{n}[H(S_1^m | W_1^m, X_2^n) - H(S_1^m | Y_1^n, W_1^m, X_2^n)], \tag{18}$$

$$= \frac{1}{n}[H(S_1^m | W_1^m) - H(S_1^m | Y_1^n, W_1^m, X_2^n)], \tag{19}$$

$$\geq \frac{1}{n}[H(S_1^m | W_1^m) - H(S_1^m | Y_1^n, W_1^m)], \tag{20}$$

$$\geq \frac{1}{b}\left[H(S_1 | W_1) - \delta(P_e^{(m,n)})\right], \tag{21}$$

where (17) follows from the Markov relation $S_1^m - X_1^n - Y_1^n$ given (X_2^n, W_1^m); (19) from the Markov relation $X_2^n - W_1^m - S_1^m$; (20) from the fact that conditioning reduces entropy; and (21) from the memoryless source assumption and from (13) which uses Fano's inequality.

On the other hand, we also have

$$
\begin{aligned}
I(X_1^n;Y_1^n|X_2^n,W_1^m) &= H(Y_1^n|X_2^n,W_1^m) - H(Y_1^n|X_1^n,X_2^n,W_1^m), && (22)\\
&= H(Y_1^n|X_2^n,W_1^m) - \sum_{i=1}^{n} H(Y_{1,i}|Y_1^{i-1},X_1^n,X_2^n,W_1^m), && (23)\\
&= H(Y_1^n|X_2^n,W_1^m) - \sum_{i=1}^{n} H(Y_{1,i}|X_{1i},X_{2i},W_1^m), && (24)\\
&\leq \sum_{i=1}^{n} H(Y_{1,i}|X_{2i},W_1^m) - \sum_{i=1}^{n} H(Y_{1,i}|X_{1i},X_{2i},W_1^m), && (25)\\
&= \sum_{i=1}^{n} I(X_{1i};Y_{1,i}|X_{2i},W_1^m), && (26)
\end{aligned}
$$

where (23) follows from the chain rule; (24) from the memoryless channel assumption; and (25) from the chain rule and the fact that conditioning reduces entropy.

For the joint mutual information we can write the following set of inequalities.

$$
\begin{aligned}
\frac{1}{n}I(X_1^n,X_2^n;Y_1^n|W_1^m) &\geq \frac{1}{n}I(S_1^m,S_2^m;Y_1^n|W_1^m), && (27)\\
&= \frac{1}{n}[H(S_1^m,S_2^m|W_1^m) - H(S_1^m,S_2^m|Y_1^n,W_1^m)], && (28)\\
&= \frac{1}{n}[H(S_1^m|W_1^m) + H(S_2^m|W_1^m) - H(S_1^m,S_2^m|Y_1^n,W_1^m)], && (29)\\
&\geq \frac{1}{n}[H(S_1^m|W_1^m) + H(S_2^m|W_1^m) - H(S_1^m,S_2^m|\hat{S}_1^m,\hat{S}_2^m), && (30)\\
&\geq \frac{1}{b}\left[H(S_1|W_1) + H(S_2|W_1) - \delta(P_e^{(m,n)})\right], && (31)
\end{aligned}
$$

where (27) follows from the Markov relation $(S_1^m,S_2^m) - (X_1^n,X_2^n) - Y_1^n$ given W_1^m; (29) from the Markov relation $S_2^m - W_1^m - S_1^m$; (30) from the fact that $(S_1^m,S_2^m) - (Y_1^n,W_1^m) - (\hat{S}_1^m,\hat{S}_2^m)$ form a Markov chain; and (31) from the memoryless source assumption and from (13) which uses Fano's inequality.

By following similar arguments as in (22)-(26) above, we can also show that

$$
I(X_1^n,X_2^n;Y_1^n|W_1^m) \leq \sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{1,i}|W_1^m). \qquad (32)
$$

Now, we introduce a time-sharing random variable $\bar{Q}$ independent of everything else. We have

$\bar{Q} = i$ with probability $1/n$, $i \in \{1, 2, \dots, n\}$. Then we can write

$$\begin{aligned}
\frac{1}{n} I(X_1^n; Y_1^n | X_2^n, W_1^m) &\leq \frac{1}{n} \sum_{i=1}^{n} I(X_{1i}; Y_{1,i} | X_{2i}, W_1^m), && (33)\\
&= \frac{1}{n} \sum_{i=1}^{n} I(X_{1\bar{q}}; Y_{\bar{q}} | X_{2\bar{q}}, W_1^m, \bar{Q} = i), && (34)\\
&= I(X_{1\bar{Q}}; Y_{\bar{Q}} | X_{2\bar{Q}}, W_1^m, \bar{Q}), && (35)\\
&= I(X_1; Y | X_2, Q), && (36)
\end{aligned}$$

where $X_1 \triangleq X_{1\bar{Q}}$, $X_2 \triangleq X_{2\bar{Q}}$, $Y \triangleq Y_{\bar{Q}}$, and $Q \triangleq (W_1^m, \bar{Q})$. Since S_1^m and S_2^m, and therefore X_{1i} and X_{2i}, are independent given W_1^m, for $q = (w_1^m, i)$ we have

$$\begin{aligned}
Pr\{X_1 = x_1, X_2 = x_2 | Q = q\} &= Pr\{X_{1i} = x_1, X_{2i} = x_2 | W_1^m = w_1^m, \bar{Q} = i\}\\
&= Pr\{X_{1i} = x_1 | W_1^m = w_1^m, \bar{Q} = i\} Pr\{X_{2i} = x_2 | W_1^m = w_1^m, \bar{Q} = i\}\\
&= Pr\{X_1 | Q = q\} \cdot Pr\{X_2 | Q = q\}.
\end{aligned}$$

Hence, the probability distribution is of the form given in Theorem 4.2.

On combining the inequalities above we can obtain

$$H(S_1|W_1) - \delta(P_e^{(m,n)}) \leq bI(X_1; Y | X_2, Q), \tag{37}$$

$$H(S_2|W_1) - \delta(P_e^{(m,n)}) \leq bI(X_2; Y | X_1, Q), \tag{38}$$

and

$$H(S_1|W_1) + H(S_2|W_1) - \delta(P_e^{(m,n)}) \leq bI(X_1, X_2; Y | Q). \tag{39}$$

Finally, taking the limit as $m, n \to \infty$ and letting $P_e^{(m,n)} \to 0$ leads to the conditions of the theorem. ∎

To the best of our knowledge, this result constitutes the first example in which the underlying source structure leads to the optimality of (informational) source-channel separation independent of the channel. We can also interpret this result as follows: The side information provided to the receiver satisfies a special Markov chain condition, which enables the optimality of source-channel separation. A natural question to ask at this point is whether providing some side information at the receiver can break the optimality of source-channel separation in the case of independent messages. In the next theorem, we show that this is not the case, and optimality of informational separation continues to hold.

Theorem 4.3: Consider independent sources S_1 and S_2 to be transmitted over the DM MAC with receiver side information W_1 as in Fig. 1. If the joint distribution satisfies $p(s_1, s_2, w_1) = p(s_1)p(s_2)p(w_1|s_1, s_2)$, then source-channel rate b is achievable if

$$H(S_1|S_2, W_1) \quad < \quad b \cdot I(X_1; Y_1|X_2, Q), \tag{40}$$

$$H(S_2|S_1, W_1) \quad < \quad b \cdot I(X_2; Y_1|X_1, Q), \tag{41}$$

and

$$H(S_1, S_2|W_1) \quad < \quad b \cdot I(X_1, X_2; Y_1|Q), \tag{42}$$

for some input distribution

$$p(q, x_1, x_2, y_1) = p(q)p(x_1|q)p(x_2|q)p(y_1|x_1, x_2), \tag{43}$$

with $|\mathcal{Q}| \leq 4$.

Conversely, if rate b is achievable, then (40)-(42) hold with $<$ replaced by $\leq$ for some joint distribution of the form given in (43).

Proof: The proof is given in Appendix I. ■

This theorem states that, if the two sources are independent, informational source-channel separation is optimal even if the receiver has side information given which independence of the sources no longer holds.

Next, we consider the case when the receiver side information is also provided to the transmitters. From the source coding perspective, i.e., when the underlying MAC is composed of orthogonal finite capacity links, it is known that having the side information at the transmitters would not help. However, it is not clear in general, from the source-channel rate perspective, whether providing the receiver side information to the transmitters would improve the performance.

If $S_1 - W_1 - S_2$ form a Markov chain, it is easy to see that the results in Theorem 4.2 continue to hold. Let $\tilde{S}_i = (S_i, W_1)$ be the new sources for which $\tilde{S}_1 - W_1 - \tilde{S}_2$ holds. Then, we have the same necessary and sufficient conditions as before, hence providing the receiver side information to the transmitters would not help in this setup.

Now, let S_i be two independent binary sequences, while $W_1 = S_1 \oplus S_2$, where $\oplus$ is the binary xor operation. In this setup, providing the receiver side information W_1 to the transmitters means that the transmitters can learn other transmitter's source as well, and hence can cooperate to transmit both sources. Source-channel rate b is achievable if

$$H(S_1, S_2|W_1) < bI(X_1, X_2; Y_1) \tag{44}$$

for some input distribution $p(x_1, x_2)$, and if rate b is achievable then (44) holds with $\leq$ for some $p(x_1, x_2)$. On the other hand, if W_1 is not available at the transmitters, we can find from Theorem 4.3 that source-channel rate b is achievable if (44) holds for some input distribution $p(x_1)p(x_2)$, and if rate b is achievable then (44) holds with $\leq$ for some $p(x_1)p(x_2)$. Thus, in this setup, providing receiver side information to the transmitters potentially leads to a smaller source-channel rate as this additional information may enable cooperation for the two source sequences that are independent of each other without the side information. We conclude that, as opposed to the pure lossless source coding scenario, having side information at the transmitters might improve the performance in the joint source-channel setting.

V. Compound MAC with correlated sources

Next, we consider a compound multiple access channel, in which two transmitters wish to transmit their correlated sources reliably to two receivers simultaneously [20]. The error probability of this system is given as follows.

$$\begin{aligned} P_e^{(m,n)} &\triangleq Pr\left\{ \bigcup_{k=1,2} (S_1^m, S_2^m) \neq (\hat{S}_{k,1}^m, \hat{S}_{k,2}^m) \right\} \\ &= \sum_{(s_1^m, s_2^m) \in \mathcal{S}_1^m \times \mathcal{S}_2^m} p(s_1^m, s_2^m) P\left\{ \bigcup_{k=1,2} (\hat{s}_{k,1}^m, \hat{s}_{k,2}^m) \neq (s_1^m, s_2^m) \Big| (S_1^m, S_2^m) = (s_1^m, s_2^m) \right\}. \end{aligned}$$

The capacity region of the compound MAC is shown to be the intersection of the two MAC capacity regions in [22] in the case of independent sources and no receiver side information. However, necessary and sufficient conditions for lossless transmission in the case of correlated sources are not known in general. In the next theorem, we extend the achievable rates in a MAC with correlated sources and side information in Theorem 4.1 to the compound MAC case. The achievability scheme is again based on correlation preserving mapping and can be extended to rates other than $b = 1$ as in [4].

Theorem 5.1: Consider lossless transmission of arbitrarily correlated sources (S_1, S_2) over a DM compound MAC with side information (W_1, W_2) at the receivers as in Fig. 1. Rate 1 is achievable if, for $k = 1, 2$,

$$\begin{aligned} H(S_1|S_2, W_k) &< I(X_1; Y_k|X_2, S_2, W_k, Q), \\ H(S_2|S_1, W_k) &< I(X_2; Y_k|X_1, S_1, W_k, Q), \\ H(S_1, S_2|U, W_k) &< I(X_1, X_2; Y_k|U, W_k, Q), \end{aligned}$$

and

$$H(S_1, S_2|W_k) < I(X_1, X_2; Y_k|W_k),$$

for some joint distribution of the form

$$p(q, s_1, s_2, w_1, w_2, x_1, x_2, y_1, y_2) = p(q)p(s_1, s_2, w_1, w_2)p(x_1|q, s_1)p(x_2|q, s_2)p(y_1, y_2|x_1, x_2)$$

and

$$U = f(S_1) = g(S_2)$$

is the common part of S_1 and S_2 in the sense of Gàcs and Körner.

Proof: The proof follows similarly to a proof in [4], and is thus omitted for the sake of brevity. ■

Note that, when there is side information at the receivers, finding the achievable source-channel rate for the compound MAC is not a simple extension of the capacity region in the case of independent sources. Due to different side information at the receivers, each transmitter should send a different part of its source to different receivers. Hence, in this case we can consider the compound MAC both as a combination of two MACs, and as a combination of two broadcast channels. We remark here that even in the case of a single broadcast channel with receiver side information, informational separation is not optimal, but the optimal source-channel rate can be achieved by operational separation [17].

In the next theorem we prove the achievability of source-channel rate b satisfying the given matching conditions. The achievability scheme is based on operational separation where the source and the channel codebooks are generated independently of each other. In particular, the typical source outputs are matched to the channel inputs without any explicit binning at the encoders. At the receiver, a joint source-channel decoder is used, which can be considered as a concatenation of a list decoder as the channel decoder, and a source decoder that searches among the list for the source codeword that is also jointly typical with the side information. However, there are no explicit source and channel codes that can be independently used either for compressing the sources or for independent data transmission over the underlying compound MAC. An alternative coding scheme composed of explicit source and channel coders that interact with each other is proposed in [18]. However, the channel code in this latter scheme is not the channel code for the underlying multi-user channel as well.

Theorem 5.2: Consider lossless transmission of arbitrarily correlated sources S_1 and S_2 over a DM compound MAC with side information W_1 and W_2 at the receivers. Rate b is achievable if, for $k = 1, 2$,

$$H(S_1|S_2, W_k) < bI(X_1; Y_k|X_2, Q), \tag{45}$$

$$H(S_2|S_1, W_k) < bI(X_2; Y_k|X_1, Q), \tag{46}$$

and

$$H(S_1, S_2|W_k) < bI(X_1, X_2; Y_k|Q), \tag{47}$$

for some $|\mathcal{Q}| \leq 4$ and input distribution of the form $p(q, x_1, x_2) = p(q)\ p(x_1|q)p(x_2|q)$.

Proof: Fix $\delta_k > 0$ and $\gamma_k > 0$ for $k = 1, 2$, and P_{X_1} and P_{X_2}. For $b = n/m$ and $k = 1, 2$, at transmitter k, we generate $M_k = 2^{m[H(S_k)+\epsilon/2]}$ i.i.d. length-m source codewords and i.i.d. length-n channel codewords using probability distributions P_{S_k} and P_{X_k}, respectively. These codewords are indexed and revealed to the receivers as well, and are denoted by $s_k^m(i)$ and $x_k^n(i)$ for $1 \leq i \leq M_k$.

Encoder: Each source outcome is directly mapped to a channel codeword as follows: Given a source outcome S_k^m at transmitter m, we find the smallest i_k such that $S_k^m = s_k^m(i_k)$, and transmit the codeword $x_k^n(i_k)$. An error occurs if no such i_k is found at either of the transmitters $k = 1, 2$.

Decoder: At receiver k, we find the unique pair (i_1^*, i_2^*) that simultaneously satisfies

$$(x_1^n(i_1^*), x_2^n(i_2^*), Y_k^n) \in \mathcal{T}^{(n)}_{[X_1 X_2 Y]_{\delta_k}},$$

and

$$(s_1^m(i_1^*), s_2^m(i_2^*), W_k^m) \in \mathcal{T}^{(m)}_{[S_1 S_2 W_k]_{\gamma_k}},$$

where $T_{[X]_\delta}^{(n)}$ is the set of weakly δ-typical sequences. An error is declared if the (i_1^*, i_2^*) pair is not uniquely determined.

Probability of error: We define the following error events:

$$E_1(k) = \{S_k^m \neq s_k^m(i), \forall i\}$$

$$E_2(k) = \{(s_1^m(i_1), s_2^m(i_2), W_k^m) \notin T_{[S_1 S_2 W_k]_{\gamma_k}}^{(m)}\}$$

$$E_3(k) = \{(x_1^n(i_1), x_2^n(i_2), Y_k^n) \notin T_{[X_1 X_2 Y]_{\delta_k}}^{(n)}\}$$

$$E_4(k) = \{\exists (j_1, j_2) \neq (i_1, i_2) : (s_1^m(j_1), s_2^m(j_2), W_k^m) \in T_{[S_1 S_2 W_k]_{\gamma_k}}^{(m)}\}$$

and

$$E_5(k) = \{\exists (j_1, j_2) \neq (i_1, i_2) : (x_1^n(j_1), x_2^n(j_2), Y_k^n) \in T_{[X_1 X_2 Y]_{\delta_k}}^{(n)}\}$$

Here, E_1 denotes the error event in which either of the encoders fails to find a unique source codeword in its codebook that corresponds to its current source outcome. When such a codeword can be found, $E_2(k)$ denotes the error event in which the sources S_1^m and S_2^m and the side information W_k at receiver k are not jointly typical, whereas $E_4(k)$ denotes the error event in which a source codeword pair different from the current realization is jointly typical with W_k. On the other hand, $E_3(k)$ denotes the error event in which channel codewords that match the current source realizations are not jointly typical with the channel output at receiver k, while $E_5(k)$ is the event in which some other channel codeword pair is jointly typical with Y_k^n.

Define $P_k^{(m,n)} \triangleq \Pr\{(S_1^m, S_2^m) \neq (\hat{S}_{k,1}^m, \hat{S}_{k,2}^m)\}$. Then $P_e^{(m,n)} \leq \sum_{k=1,2} P_k^{(m,n)}$. Again, from the union bound, we have

$$\begin{aligned} P_k^{(m,n)} \leq & \Pr\{E_1(k)\} + \Pr\{E_2(k)\} + \Pr\{E_3(k)\} \\ & + \sum_{\substack{j_1 \neq i_1, \\ j_2 = i_2}} \Pr\left\{(s_1^m(j_1), s_2^m(j_2), W_k^m) \in T_{[S_1, S_2, W_k]_{\gamma_k}}^{(m)}\right\} \Pr\left\{(x_1^n(j_1), x_2^n(j_2), Y_k^n) \in T_{[X_1, X_2, Y_k]_{\delta_k}}^{(n)}\right\} \\ & + \sum_{\substack{j_1 = i_1, \\ j_2 \neq i_2}} \Pr\{\cdot\} + \sum_{\substack{j_1 \neq i_1, \\ j_2 \neq i_2}} \Pr\{\cdot\}. \end{aligned} \tag{48}$$

In [17] it is shown that, for any $\lambda > 0$ and sufficiently large m

$$\begin{aligned} \Pr\{E_1(k)\} &= (1 - \Pr\{S_k^m = s_k^m(1)\})^{M_k} \\ &\leq \exp^{-2^{-n[H(S_k)+6\lambda]} M_k} \\ &= \exp^{-2^{n[\frac{\epsilon}{2} - 6\lambda]}}. \end{aligned} \tag{49}$$

We choose $\lambda < \frac{\epsilon}{12}$, and obtain $\Pr\{E_1\} \to 0$ as $m \to \infty$.

Similarly, we can also prove that $\Pr(E_i(k)) \to 0$ for $i = 2, 3$ and $k = 1, 2$ as $m, n \to \infty$ using standard techniques. We can also obtain

$$\sum_{\substack{j_1 \neq i_1, \\ j_2 = i_2}} \Pr\left\{(s_1^m(j_1), s_2^m(j_2), W_k^m) \in T^{(m)}_{[S_1,S_2,W_k]_{\gamma_k}}\right\} \Pr\left\{(x_1^n(j_1), x_2^n(j_2), Y_k^n) \in T^{(n)}_{[X_1 X_2 Y_k]_{\delta_k}}\right\}$$

$$\leq 2^{m[H(S_1)+\frac{\epsilon}{2}]-m[I(S_1;S_2,W_k)-\lambda]-n[I(X_1;Y_k|X_2)-\lambda]} \tag{50}$$

$$= 2^{-m[H(S_1|S_2,W_k)-bI(X_1;Y_k|X_2)-(b+1)\lambda-\frac{\epsilon}{2}]}$$

$$= 2^{-m[\frac{\epsilon}{2}-(b+1)\lambda]} \tag{51}$$

where in (50) we used (1) and (2); and (51) holds if the conditions in the theorem hold.

A similar bound can be found for the second summation in (48). For the third one, we have the following bound.

$$\sum_{\substack{j_1 \neq i_1, \\ j_2 \neq i_2}} \Pr\left\{(s_1^m(j_1), s_2^m(j_2), W_k^m) \in T^{(m)}_{[S_1,S_2,W_k]_{\gamma_k}}\right\} \Pr\left\{(x_1^n(j_1), x_2^n(j_2), Y_k^n) \in T^{(n)}_{[X_1,X_2,Y]_{\delta_k}}\right\}$$

$$\leq 2^{m[H(S_1)+\epsilon/2]+m[H(S_2)+\epsilon/2]} 2^{-m[I(S_1;S_2,W_k)+I(S_2;S_1,W_k)-I(S_1;S_2|W_k)]-\lambda]} 2^{-n[I(X_1,X_2;Y_k)-\lambda]} \tag{52}$$

$$\leq 2^{-m[H(S_1|S_2,W_k)+H(S_2|S_1,W_k)-bI(X_1,X_2;Y_k)-(b+1)\lambda-\epsilon]}$$

$$= 2^{-m[\epsilon-(b+1)\lambda]}, \tag{53}$$

where (52) follows from (1) and (3); and (53) holds if the conditions in the theorem hold.

Choosing $\lambda < \frac{\epsilon}{2(b+1)}$, we can make sure that all these three terms also vanish as $m, n \to \infty$. Any rate pair in the convex hull can be achieved by time sharing, hence the time-sharing random variable Q. The cardinality bound on Q follows from the classical arguments. ■

We next prove that the conditions in Theorem 5.2 are also necessary to achieve a source-channel rate of b for some special settings. We first consider the case in which (S_1, W_2) is independent of (S_2, W_1). This might model a scenario in which Rx_1 (Rx_2) and Tx_2 (Tx_1) are located close to each other, thus having correlated observations, while the two transmitters are far away from each other (see Fig. 2).

Theorem 5.3: Consider lossless transmission of arbitrarily correlated sources S_1 and S_2 over a DM compound MAC with side information W_1 and W_2, where (S_1, W_2) is independent of (S_2, W_1). Separation (in operational sense) is optimal for this setup, and rate b is achievable if,

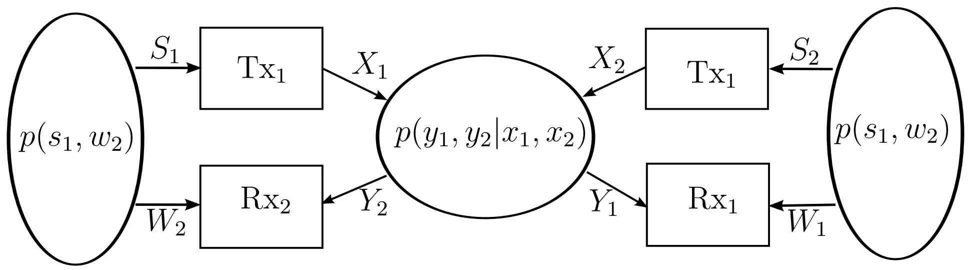

Fig. 2. Compound multiple access channel in which the transmitter 1 (2) and receiver 2 (1) are located close to each other, and hence have correlated observations, independent of the other pair, i.e., (S_1, W_2) is independent of (S_2, W_1).

for $(k, m) \in \{(1,2),(2,1)\}$

$$H(S_k) < bI(X_k; Y_k|X_m, Q), \tag{54}$$

$$H(S_m|W_k) < bI(X_m; Y_k|X_k, Q), \tag{55}$$

and

$$H(S_k) + H(S_m|W_k) < bI(X_k, X_m; Y_k|Q), \tag{56}$$

for some $|\mathcal{Q}| \leq 4$ and input distribution of the form

$$p(q, x_1, x_2) = p(q)p(x_1|q)p(x_2|q). \tag{57}$$

Conversely, if rate b is achievable, then (54)-(56) hold with $<$ replaced by $\leq$ for an input probability distribution of the form given in (57).

Proof: Achievability follows from Theorem 5.2, and the converse proof is given in Appendix II. ■

Next, we consider the case in which there is no multiple access interference at the receivers (see Fig. 3). We let $Y_k = (Y_{1,k}, Y_{2,k})$ $k = 1, 2$, where the memoryless channel is characterized by

$$p(y_{1,1}, y_{2,1}, y_{1,2}, y_{2,2}|x_1, x_2) = p(y_{1,1}, y_{1,2}|x_1)p(y_{2,1}, y_{2,2}|x_2). \tag{58}$$

On the other hand, we allow arbitrary correlation among the sources and the side information. However, since there is no multiple access interference, using the source correlation to create correlated channel codewords does not enlarge the rate region of the channel. We also remark

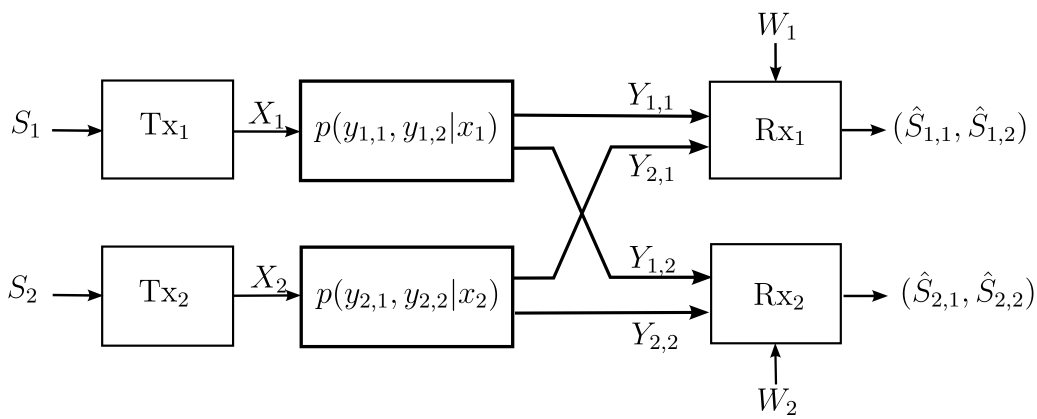

Fig. 3. Compound multiple access channel with correlated sources and correlated side information with no multiple access interference.

that, this model is not equivalent to two independent broadcast channels with side information. The two encoders interact with each other through the correlation among their sources.

Theorem 5.4: Consider lossless transmission of arbitrarily correlated sources S_1 and S_2 over a DM compound MAC with no multiple access interference characterized by (58) and receiver side information W_1 and W_2 (see Fig. 3). Separation (in operational sense) is optimal for this setup, and rate b is achievable if, for $(k, m) = \{(1, 2), (2, 1)\}$

$$H(S_k|S_m, W_k) \quad < \quad bI(X_k; Y_{k,k}), \tag{59}$$

$$H(S_m|S_k, W_k) \quad < \quad bI(X_m; Y_{m,k}), \tag{60}$$

and

$$H(S_k, S_m|W_k) \quad < \quad b[I(X_k; Y_{k,k}) + I(X_m; Y_{m,k})], \tag{61}$$

for an input distribution of the form

$$p(q, x_1, x_2) = p(q)p(x_1|q)p(x_2|q). \tag{62}$$

Conversely, if rate b is achievable, then (54)-(56) hold with $<$ replaced by $\leq$ for an input probability distribution of the form given in (57).

Proof: The achievability follows from Theorem 5.2 by letting Q be constant and taking into consideration the characteristics of the channel, where $(X_1, Y_{1,1}, Y_{1,2})$ is independent of $(X_2, Y_{2,1}, Y_{2,2})$. The converse can be proven similarly to Theorem 5.3, and will be omitted for the sake of brevity. ■

Note that the model considered in Theorem 5.4 is a generalization of the model in [23] (which is a special case of the more general network studied in [14]) to more than one receiver. Theorem 5.4 considers correlated receiver side information which can be incorporated into the model of [23] by considering an additional transmitter sending this side information over an infinite capacity link. In this case, using [23], we observe that informational source-channel separation is optimal. However, Theorem 5.4 argues that this is no longer true when the number of sink nodes is greater than one even when there is no receiver side information.

The model in Theorem 5.4 is also considered in [24] in the special case of no side information at the receivers. In the achievability scheme of [24], transmitters first randomly bin their correlated sources, and then match the bins to channel codewords. Theorem 5.4 shows that we can achieve the same optimal performance without explicit binning even in the case of correlated receiver side information.

In both Theorem 5.3 and Theorem 5.4, we provide the optimal source-channel matching conditions for lossless transmission. While general matching conditions are not known for compound MACs, the reason we are able to resolve the problem in these two cases is the lack of multiple access interference from users with correlated sources. In the first setup the two sources are independent, hence it is not possible to generate correlated channel inputs, while in the second setup, there is no multiple access interference, and thus there is no need to generate correlated channel inputs.

Finally, we consider the special case in which the two receivers share common side information, i.e., $W_1 = W_2 = W$. Moreover, for this common side information, similar to Theorem 4.2, $S_1 - W - S_2$ form a Markov chain. For example this models the scenario in which the two receivers are close to each other, hence they have the same side information. The following theorem proves informational separation under these conditions.

Theorem 5.5: Consider lossless transmission of correlated sources S_1 and S_2 over a DM compound MAC with common receiver side information $W_1 = W_2 = W$ satisfying $S_1 - W - S_2$. Separation (in informational sense) is optimal in this setup, and rate b is achievable if, for $k = 1$ and 2,

$$
\begin{aligned}
H(S_1|W) &< b \cdot I(X_1; Y_k|X_2, Q), \qquad (63)\\
H(S_2|W) &< b \cdot I(X_2; Y_k|X_1, Q),
\end{aligned}
$$

and

$$H(S_1|W) + H(S_2|W) \;\; < \;\; b \cdot I(X_1, X_2; Y_k|Q),$$

for some joint distribution $p(q, x_1, x_2, y) = p(q)p(x_1|q)\; p(x_2|q)p(y|x_1, x_2)$, with $|\mathcal{Q}| \leq 4$.

Conversely, if rate b is achievable, then (63)-(64) hold with $<$ replaced by $\leq$ for an input probability distribution of the form given above.

Proof: The achievability follows from informational source-channel separation, i.e, Slepian-Wolf compression conditioned on the receiver side information followed by an optimal compound MAC coding. The proof of the converse follows similarly to the proof of Theorem 4.2, and is omitted for brevity. ■

VI. Interference channel with correlated sources

In this section, we consider the interference channel (IC) with correlated sources and side information. In the IC each transmitter wishes to communicate only with its corresponding receiver, while the two simultaneous transmissions interfere with each other. Even when the sources and the side information are all independent, the capacity region of the IC is in general not known. The best achievable scheme is given in [25]. The capacity region can be characterized in the strong interference case [34], [35], where it coincides with the capacity region of the compound multiple access channel, i.e., it is optimal for the receivers to decode both messages. Interference channels has gained recent interest due to its practical value in cellular and cognitive radio systems. See [26] - [33] and references therein for recent results relating to the capacity region of various interference channel scenarios.

For encoders $f_i^{(m,n)}$ and decoders $g_i^{(m,n)}$, the probability of error for the interference channel is given as

$$\begin{aligned} P_e^{(m,n)} \;\; &\triangleq \;\; Pr\left\{ \bigcup_{k=1,2} S_k^m \neq \hat{S}_{k,k}^m \right\} \\ &= \sum_{(s_1^m, s_2^m) \in \mathcal{S}_1^m \times \mathcal{S}_2^m} p(s_1^m, s_2^m) P\left\{ \bigcup_{k=1,2} \hat{s}_{k,k}^m \neq s_k^m \Big| (S_1^m, S_2^m) = (s_1^m, s_2^m) \right\}. \end{aligned}$$

In the case of correlated sources and receiver side information, sufficient conditions for the achievability of source-channel rate 1 over the compound MAC given in Theorem 4.1 serve as sufficient conditions for the IC as well, since we can constrain both receivers to obtain lossless reconstruction of both sources. Our goal here is to characterize the conditions under which we can provide a converse and achieve either informational or operational separation similar to the results of Section V.

We first note that the sufficient conditions of Theorem 5.2 for achievability of rate b using operational separation in compound MACs hold for ICs as well. In order to extend the necessary conditions of Theorem 5.3 and Theorem 5.5 to ICs, we will define the 'strong source-channel interference' conditions. Note that the interference channel version of Theorem 5.4 where there is no multiple access interference is trivial since the two transmissions do not interfere with each other.

The regular strong interference conditions given in [34] correspond to the case in which, for all input distributions at transmitter Tx_1, the rate of information flow to receiver 2 is higher than the information flow to the intended receiver Rx_1. A similar condition holds for transmitter Tx_2 as well. Hence there is no rate loss if both receivers decode the messages of both transmitters. Consequently, under strong interference conditions, the capacity region of the IC is equivalent to the capacity region of the compound MAC. However, in the joint source-channel coding scenario, the receivers have access to correlated side information. Thus while calculating the total rate of information flow to a particular receiver, we should not only consider the information flow through the channel, but also the mutual information that already exists between the source and the receiver side information.

We first focus on the scenario of Theorem 5.3 in which the two sources are independent while the side information W_1 is correlated with source S_2, and the side information W_2 is correlated with source S_1.

Definition 6.1: For the interference channel in Fig. 2 in which (S_1, W_2) is independent of (S_2, W_1), we say that the *strong source-channel interference conditions* are satisfied if

$$b \cdot I(X_1; Y_1|X_2) \leq b \cdot I(X_1; Y_2|X_2) + I(S_1; W_2), \tag{64}$$

and

$$b \cdot I(X_2; Y_2|X_1) \leq b \cdot I(X_2; Y_1|X_1) + I(S_2; W_1), \tag{65}$$

for all distributions of the form $p(w_1, w_2, s_1, s_2, x_1, x_2) = p(w_1, w_2, s_1, s_2)p(x_1|s_1)p(x_2|s_2)$.

For an IC satisfying these conditions, we next prove the following theorem.

Theorem 6.1: Consider lossless transmission of S_1 and S_2 over a DM IC with side information W_1 and W_2, where (S_1, W_2) is independent of (S_2, W_1) as in Fig. 2. Under the strong source-channel interference conditions of Definition 6.1, separation (in operational sense) is optimal and source-channel rate b is achievable if the conditions (45)-(47) in Theorem 5.2 hold. Conversely, if rate b is achievable, then the conditions in Theorem 5.2 hold with $<$ replaced by $\leq$.

Before we proceed with the proof of the theorem, we first prove the following lemma.

Lemma 6.2: If (S_1, W_2) is independent of (S_2, W_1) and the strong source-channel interference conditions (64)-(65) hold, then we have

$$I(X_2^n; Y_2^n | X_1^n) \;\leq\; I(X_2^n; Y_1^n | X_1^n) + I(S_2^m; W_1^m), \tag{66}$$

and

$$I(X_1^n; Y_1^n | X_2^n) \;\leq\; I(X_1^n; Y_2^n | X_2^n) + I(S_1^m; W_2^m), \tag{67}$$

for all m and n satisfying $n/m = b$.

Proof: To prove the lemma, we follow the techniques in [35]. Condition (65) implies

$$I(X_2; Y_2 | X_1, U) - I(X_2; Y_1 | X_1, U) \leq \frac{1}{b} I(S_2; W_1) \tag{68}$$

for all U satisfying $U - (X_1, X_2) - (Y_1, Y_2)$.

Then as in [35], we can obtain

$$\begin{aligned}
I(X_2^n; Y_2^n | X_1^n) - I(X_2^n; Y_1^n | X_1^n) =& I(X_{2n}; Y_{2n} | X_1^n, Y_2^{n-1}) - I(X_{2n}; Y_{1n} | X_1^n, Y_2^{n-1}) \\
& + I(X_2^{n-1}; Y_2^{n-1} | X_1^n, Y_{1n}) - I(X_2^{n-1}; Y_1^{n-1} | X_1^n, Y_{1n}) \\
=& I(X_{2n}; Y_{2n} | X_{1n}) - I(X_{2n}; Y_{1n} | X_{1n}) \\
& + I(X_2^{n-1}; Y_2^{n-1} | X_1^{n-1}) - I(X_2^{n-1}; Y_1^{n-1} | X_1^{n-1}) \\
=& \sum_{i=1}^{n} [I(X_{2i}; Y_{2i} | X_{1i}) - I(X_{2i}; Y_{1i} | X_{1i})].
\end{aligned}$$

Using the hypothesis (65) of the theorem, we obtain

$$\begin{aligned}
I(X_2^n; Y_2^n | X_1^n) - I(X_2^n; Y_1^n | X_1^n) &\leq \frac{n}{b} I(S_2; W_1) \\
&= I(S_2^m; W_1^m).
\end{aligned}$$

Eqn. (67) follows similarly. ∎

Proof: (of Theorem 6.1) Achievability follows by having each receiver decode both S_1 and S_2, and then using Theorem 5.1. We next prove the converse. From (81)-(84), we have

$$\frac{1}{n}I(X_1^n;Y_1^n|X_2^n) \geq \frac{1}{b}\left[H(S_1) - \delta(P_e^{(m,n)})\right]. \tag{69}$$

We can also obtain

$$\begin{aligned}
\frac{1}{n}I(X_1^n;Y_2^n|X_2^n) &\geq \frac{1}{n}[I(X_1^n;Y_1^n|X_2^n) - I(S_1^m;W_2^m)], &(70)\\
&= \frac{1}{b}[H(S_1) - \delta(P_e^{(m,n)})] - \frac{1}{n}I(S_1^m;W_2^m), &(71)\\
&= \frac{1}{b}[H(S_1|W_2) - \delta(P_e^{(m,n)})], &(72)
\end{aligned}$$

in which (70) follows from (67), and (71) from (69).

Finally for the joint mutual information, we have

$$\begin{aligned}
\frac{1}{n}I(X_1^n,X_2^n;Y_1^n) =&\frac{1}{n}[I(X_1^n;Y_1^n) + I(X_2^n;Y_1^n|X_1^n)],\\
\geq&\frac{1}{n}[I(S_1^m;Y_1^n) + I(X_2^n;Y_2^n|X_1^n) - I(S_2^m;W_1^m)], &(73)\\
\geq&\frac{1}{n}[I(S_1^m;Y_1^n) + I(S_2^m;Y_2^n|X_1^n) - I(S_2^m;W_1^m)], &(74)\\
=&\frac{1}{n}[H(S_1^m) - H(S_1^m|Y_1^n) + H(S_2^m|X_1^n) - H(S_2^m|Y_2^n,X_1^n)\\
&\quad + H(S_2^m|W_1^m) - H(S_2^m)],\\
\geq&\frac{1}{n}[H(S_1^m) - H(S_1^m|Y_1^n) - H(S_2^m|Y_2^n) + H(S_2^m|W_1^m)], &(75)\\
=&\frac{1}{n}[H(S_1^m) - H(S_1^m|Y_1^n,W_1^m) - H(S_2^m|Y_2^n,W_2^m) + H(S_2^m|W_1^m)], &(76)\\
\geq&\frac{1}{b}[H(S_1) + H(S_2|W_1) - 2\delta(P_e^{(m,n)})], &(77)
\end{aligned}$$

for any $\epsilon > 0$ and large enough m and n, where (73) follows from the data processing inequality and (66); (74) follows from the data processing inequality since $S_2^m - X_2^n - Y_2^n$ form a Markov chain given X_1^n; (75) follows from the independence of X_1^n and S_2^m and the fact that conditioning reduces entropy; (76) follows from the independence of (S_1^m, W_2^m) and (S_2^m, W_1^m); and (77) follows from Fano's inequality.

The rest of the proof closely resembles the proof of Theorem 5.3. ■

Next, we consider the IC version of the case in Theorem 5.5, in which the two receivers have access to the same side information W and with this side information the sources are independent. While we still have correlation between the sources and the common receiver side information, the amount of mutual information arising from this correlation is equivalent at both receivers since $W_1 = W_2$. This suggests that the usual strong interference channel conditions suffice to obtain the converse result. We have the following theorem for this case.

Theorem 6.3: Consider lossless transmission of correlated sources S_1 and S_2 over the strong IC with common receiver side information $W_1 = W_2 = W$ satisfying $S_1 - W - S_2$. Separation (in informational sense) is optimal in this setup, and source-channel rate b is achievable if and only if the conditions in Theorem 5.5 hold.

Proof: The proof follows from arguments similar to those in the proof of Theorem 5.5 and results in [19], where we incorporate the strong source-channel interference conditions. ■

VII. Conclusions

We have considered source and channel coding over multi-user channels with correlated receiver side information. Due to the lack of a general source-channel separation theorem for multi-user channels, optimal performance in general requires joint source-channel coding. Moreover, we do not know the exact source-channel matching conditions even for simple channels. In this paper, we have characterized the necessary and sufficient conditions for lossless transmission over various fundamental multi-user channels, such as multiple access, compound multiple access and interference channels with certain source-channel structures. In particular, we have considered transmitting correlated sources over MAC with receiver side information given which the sources are independent, and transmitting independent sources over MAC with receiver side information given which the sources are correlated. For the compound MAC, we have provided an achievability result, which has been shown to be tight i) when each source is correlated with only one of the side information and the two source-side information pairs are independent of each other; and ii) when the sources and the side information are arbitrarily correlated but there is no multiple access interference at the receivers, iii) when the sources are correlated and the receivers have access to the same side information given which the two sources are independent. We have then showed that for the cases (i) and (iii), the conditions provided for compound MAC are also necessary for interference channels under some strong source-channel conditions.

For the cases analyzed in this paper, we have proven the optimality of designing source and channel codes that are statistically independent of each other, hence resulting in a modular system design without losing the end-to-end optimality. These results will help to increase our understanding of the fundamental limits of joint source-channel coding in more complicated networks, and hopefully will lead to improved design principles for practical implementations.

APPENDIX I
PROOF OF THEOREM 4.3

Proof: The achievability again follows from separate source and channel coding. We first use Slepian-Wolf compression of the sources conditioned on the receiver side information, then transmit the compressed messages using an optimal multiple access channel code.

For the converse, we use Fano's inequality given in (13) and (16). We have

$$\begin{aligned}
\frac{1}{n}I(X_1^n;Y_1^n|X_2^n) &\geq \frac{1}{n}I(S_1^m;Y_1^n|X_2^n), && (78)\\
&= \frac{1}{n}I(S_1^m,W_1^m;Y_1^n|X_2^n), && (79)\\
&\geq \frac{1}{n}I(S_1^m;Y_1^n|X_2^n,W_1^m), \\
&\geq \frac{1}{n}[H(S_1^m|S_2^m,W_1^m) - m\delta(P_e^{(m,n)})], && (80)\\
&\geq \frac{1}{b}[H(S_1|S_2,W_1) - \delta(P_e^{(m,n)})],
\end{aligned}$$

where (78) follows from the Markov relation $S_1^m - X_1^n - Y_1^n$ given X_2^n; (79) from the Markov relation $W_1^m - (X_2^n, S_1^m) - Y_1^n$; and (80) from Fano's inequality (16).

We also have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} I(X_{1i};Y_{1,i}|X_{2i}) &\geq \frac{1}{n}I(X_1^n,X_2^n;Y_1^n)\\
&\geq \frac{1}{b}[H(S_1|S_2,W_1) - \delta(P_e^{(m,n)})].
\end{aligned}$$

Similarly, we have

$$\frac{1}{n}\sum_{i=1}^{n} I(X_{2i};Y_{1,i}|X_{1i}) \geq \frac{1}{b}[H(S_2|S_1,W_1) - \delta(P_e^{(m,n)})],$$

and

$$\frac{1}{n}\sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{1,i}) \geq \frac{1}{b}[H(S_1,S_2|W_1) - \delta(P_e^{(m,n)})].$$

As usual, we let $P_e^{(m,n)} \rightarrow 0$, and introduce the time sharing random variable Q uniformly distributed over $\{1, 2, \ldots, n\}$ and independent of all the other random variables. Then we define $X_1 \triangleq X_{1Q}$, $X_2 \triangleq X_{2Q}$ and $Y_1 \triangleq Y_{1Q}$. Note that $Pr\{X_1 = x_1, X_2 = x_2 | Q = q\} = Pr\{X_1 | Q = q\} \cdot Pr\{X_2 | Q = q\}$ since the two sources, and hence the channel codewords, are independent of each other conditioned on Q. Thus, we obtain (40)-(42) for a joint distribution of the form (43). ■

APPENDIX II
PROOF OF THEOREM 5.3

We have

$$\begin{aligned}
\frac{1}{n} I(X_1^n; Y_1^n | X_2^n) \geq & \frac{1}{n} I(S_1^m; Y_1^n | X_2^n), && (81)\\
= & \frac{1}{n} [H(S_1^m | X_2^n) - H(S_1^m | Y_1^n, X_2^n)], && (82)\\
\geq & \frac{1}{n} [H(S_1^m) - H(S_1^m | Y_1^n)], && (83)\\
\geq & \frac{1}{b} \left[H(S_1) - \delta(P_e^{(m,n)}) \right], && (84)
\end{aligned}$$

for any $\epsilon > 0$ and sufficiently large m and n, where (81) follows from the conditional data processing inequality since $S_1^m - X_1^n - Y_1^n$ forms a Markov chain given X_2^n; (83) from the independence of S_1^m and X_2^n and the fact that conditioning reduces entropy; and (84) from the memoryless source assumption, and from Fano's inequality.

For the joint mutual information, we can write the following set of inequalities:

$$\begin{aligned}
& \frac{1}{n} I(X_1^n, X_2^n; Y_1^n) \geq \frac{1}{n} I(S_1^m, S_2^m; Y_1^n), && (85)\\
& = \frac{1}{n} I(S_1^m, S_2^m, W_1^m; Y_1^n), && (86)\\
& \geq \frac{1}{n} I(S_1^m, S_2^m; Y_1^n | W_1^m), && (87)\\
& = \frac{1}{n} [H(S_1^m, S_2^m | W_1^m) - H(S_1^m, S_2^m | Y_1^n, W_1^m)], \\
& = \frac{1}{n} [H(S_1^m) + H(S_2^m | W_1^m) - H(S_1^m, S_2^m | Y_1^n, W_1^m)], && (88)\\
& \geq \frac{1}{b} \left[H(S_1) + H(S_2 | W_1) - \delta(P_e^{(m,n)}) \right], && (89)
\end{aligned}$$

for any $\epsilon > 0$ and sufficiently large m and n, where (85) follows from the data processing inequality since $(S_1^m, S_2^m) - (X_1^n, X_2^n) - Y_1^n$ form a Markov chain; (86) from the Markov relation $W_1^m - (S_1^m, S_2^m) - Y_1^n$; (87) from the chain rule and the non-negativity of the mutual information; (88) from the independence of S_1^m and (S_2^m, W_1^m); and (89) from the memoryless source assumption and Fano's inequality.

It is also possible to show that

$$\sum_{i=1}^{n} I(X_{1i}; Y_{1i}|X_{2i}) \geq I(X_1^n; Y_1^n|X_2^n), \tag{90}$$

and similarly for other mutual information terms. Then, using the above set of inequalities and letting $P_e^{(m,n)} \to 0$, we obtain

$$\begin{aligned} \frac{1}{b}H(S_1) &\leq \frac{1}{n}\sum_{i=1}^{n} I(X_{1i}; Y_{1i}|X_{2i}), \\ \frac{1}{b}H(S_2|W_1) &\leq \frac{1}{n}\sum_{i=1}^{n} I(X_{2i}; Y_{1i}|X_{1i}), \end{aligned}$$

and

$$\frac{1}{b}(H(S_1) + H(S_2|W_1)) \leq \frac{1}{n}\sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_{1i}),$$

for any product distribution on $\mathcal{X}_1 \times \mathcal{X}_2$. We can write similar expressions for the second receiver as well. Then the necessity of the conditions of Theorem 5.2 can be argued simply by inserting the time-sharing random variable Q.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, July 1948.

[2] S. Vembu, S. Verdú and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 44-54, Jan. 1995.

[3] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Satist. Probability*, vol. 1, 1961, pp. 611-644.

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

[5] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471-480, Jul. 1973.

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

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

[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. 29, no.3, pp.396412, May 1983.

[9] T. S. Han and M. H. M. Costa, “Broadcast channels with arbitrarily correlated sources,” *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 641-650, Sep. 1987.

[10] M. Salehi and E. Kurtas, “Interference channels with correlated sources,” *Proc. IEEE International Symposium on Information Theory*, San Antonio, TX, Jan. 1993.

[11] S. Choi and S. S. Pradhan, “A graph-based framework for transmission of correlated sources over broadcast channels,” *IEEE Trans. Inf. Theory*, 2008.

[12] 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.

[13] W. Kang and S. Ulukus, “A new data processing inequality and its applications in distributed source and channel coding,” submitted to *IEEE Trans. Inf. Theory*.

[14] T. S. Han, “Slepian-Wolf-Cover theorem for a network of channels,” *Information and Control*, vol. 47, no. 1, pp. 67-83, 1980.

[15] K. de Bruyn, V. V. Prelov and E. van der Meulen, “Reliable transmission of two correlated sources over an asymmetric multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 716-718, Sep. 1987.

[16] M. Effros, M. Médard, T. Ho, S. Ray, D. Karger and R. Koetter, “Linear network codes: A Unified Framework for Source Channel, and Network Coding,” in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 66, American Mathematical Society, 2004.

[17] E. Tuncel, “Slepian-Wolf coding over broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469-1482, April 2006.

[18] D. Gündüz and E. Erkip, “Reliable cooperative source transmission with side information,” *Proc. IEEE Information Theory Workshop*, Bergen, Norway, July 2007.

[19] D. Gündüz and E. Erkip, “Lossless transmission of correlated sources over a multiple access channel with side information,” *Proc. Data Compression Conference*, Snowbird, UT, March 2007.

[20] D. Gündüz and E. Erkip, “Interference channel and compound MAC with correlated sources and receiver side information,” *Proc. IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[21] P. Gács and J. Körner, “Common information is far less than mutual information,” *Problems of Control and Information Theory*, vol. 2, no. 2, pp. 149-162, 1973.

[22] R. Ahlswede, “The capacity region of a channel with two senders and two receivers,” *Annals of Probability*, vol. 2, no. 5, pp. 805-814, 1974.

[23] J. Barros and S. D. Servetto, “Network information flow with correlated sources,” *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 155-170, Jan. 2006.

[24] T. P. Coleman, E. Martinian and E. Ordentlich, “Joint source-channel decoding for transmitting correlated sources over broadcast networks,” submitted to *IEEE Trans. Inf. Theory*.

[25] T. S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 27, no.1, pp. 49-60, Jan. 1981.

[26] W. Wu, S. Vishwanath, and A. Arapostathis, “On the capacity of Gaussian weak interference channels with degraded message sets,” Proc. 40th Annual Conference on Information Sciences and Systems, Princeton, NJ, Mar. 2006.

[27] H. Chong, M. Motani, H. Garg and H. El Gamal, "On the Han-Kobayashi region for the interference channel," *submitted to IEEE Trans. on Inf. Theory*, 2006.

[28] I. Maric, R. D. Yates and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. on Inf. Theory*, vol. 53, no. 10, pp. 3536-3548, Oct. 2007.

[29] R. Etkin, D. Tse, H. Wang, "Gaussian interference channel capacity to within one bit," submitted to IEEE Trans. on Information Theory, Feb. 2007.

[30] I. E. Telatar and D. Tse, "Bounds on the capacity region for a class of interference channels," *Proc. IEEE International Symposium on Information Theory*, Nice, France, July 2007.

[31] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *submitted to the IEEE Trans. Inform. Theory.*

[32] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *submitted to IEEE Trans. Inform. Theory.*

[33] V. S. Annapureddy and V. Veeravalli, "Sum capacity of the Gaussian interference channel in the low interference regime," submitted to *submitted to IEEE Trans. Inform. Theory.*

[34] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786-788, Nov. 1981.

[35] M. H. M. Costa and A. A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. on Information Theory*, vol. 33, no. 5, pp. 710-711, Sep. 1987.