On Necessary Conditions for Multiple-Access-Relay Channels with Correlated Sources

Yonathan Murin, Ron Dabora
Department of Electrical and Computer Engineering
Ben-Gurion University, Israel
Email: moriny@ee.bgu.ac.il, ron@ee.bgu.ac.il

Deniz Gündüz
Dept. of Electrical and Electronic Engineering
Imperial College London, United Kingdom
Email: d.gunduz@imperial.ac.uk

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. The optimal joint source-channel coding scheme for transmission of correlated sources over multiple-access-relay channels (MARCs) is an open problem. Here, this problem is studied in the presence of arbitrarily correlated side information at both the relay and destination. Since each transmitter observes a single source sequence, the admissible joint distributions of the sources and channel inputs must satisfy a Markov relationship which restricts the statistical dependence of the channel inputs. This observation is used jointly with the new data processing inequality proposed by [Kang and Ulukus, 2011] to derive two new sets of single-letter necessary conditions. These new conditions are shown to be at least as tight as the previously known ones, and strictly tighter than the cut-set bound.**

I. Introduction

The multiple-access-relay channel (MARC) is a multiuser network in which several sources communicate with a single destination, assisted by a common relay [1]. We study lossless transmission of arbitrarily correlated sources over MARCs, assuming that both the relay and the destination have access to side information that is correlated with the sources. This model may represent cooperative transmission of correlated observations to an access point over wireless sensor networks.

The realization of a random source can be reliably transmitted over a memoryless point-to-point (PtP) channel if its entropy is less than the channel capacity [2]. Conversely, if the source entropy is larger than the channel capacity, the source cannot be reliably transmitted over the channel. This result tells us that the optimal end-to-end performance can be achieved by separate source-channel coding, that is, by first compressing the source at a rate equal to its entropy, and then transmitting the compressed bits over the channel using a capacity-achieving channel code. It is well-known that the optimality of separate design does not generalize to multi-user networks [3], and in general, optimal performance requires a joint design of the source and channel codes. An important technique for joint source-channel coding (JSCC) is the correlation preserving mapping (CPM) technique, introduced in [3] for the transmission of correlated sources over discrete, memoryless (DM) multiple access channels (MACs). In [6] and [7] we applied the CPM technique to DM MARCs and proposed four new joint source-channel coding (JSCC) schemes for reliable transmission of correlated sources over DM MARCs. Each of these schemes implements a different combination of Slepian-Wolf source coding and the CPM technique. It is shown in [7] that these four schemes constitute four different sets of sufficient conditions, not necessarily superior to each other.

Here we focus on the necessary conditions for reliable transmission of correlated sources over DM MARCs. Such conditions were previously derived in [8] and [9]. Observe that the admissible joint distributions of the sources and the respective channel inputs for the MARC (and for the MAC), must satisfy a Markov relationship which reflects the fact that the channel inputs at the transmitters are correlated *only via the correlation of the sources*. This fact is not accounted for in the conditions in [9]. On the other hand, while [8] established necessary conditions which account for the above constraint, these conditions are based on n-letter mutual information expressions, and therefore they are not computable. n-letter necessary conditions for transmission of correlated sources over MAC were originally derived in [3]. In [4], Kang and Ulukus used spectral analysis to introduce a new set of single-letter necessary conditions for reliable transmission of correlated sources over MACs, which both accounts for the Markov relation and leads to computable single-letter conditions.

Main Contributions

We derive three new sets of single-letter necessary conditions for reliable transmission of correlated sources over DM MARCs. The first set is in the spirit of a MAC bound for the classic relay channel, while the other two are in line with the broadcast bound [10, Ch. 16]. Similarly to [4], the proposed sets take into account the Markov relation between the sources and the channel inputs, which tightens the conditions with respect to the cut-set bound. The proposed conditions are shown to be at least as tight as the ones derived in [9], and not equivalent to each other. Furthermore, through a numerical example, we show that the proposed necessary conditions are strictly tighter than the cut-set bound [10, Ch. 18.1].

This paper is organized as follows: The system model and notations are introduced in Section II. In Section III we give some preliminaries, based on [4]. The new sets of necessary conditions are presented in Section IV. A numerical example is given in Section V, and the paper is concluded in Section VI.

II. Notations and Model

In this work, we denote discrete random variables (RVs) with capital letters, their realizations with lower case letters, and their alphabets by the respective calligraphic letters. We use $|\mathcal{X}|$ to denote the cardinality of a finite, discrete set $\mathcal{X}$, and

This work was partially supported by the European Commission's Marie Curie IRG Fellowship PIRG05-GA-2009-246657 under the Seventh Framework Programme, and by the Israel Science Foundation under grant 396/11.

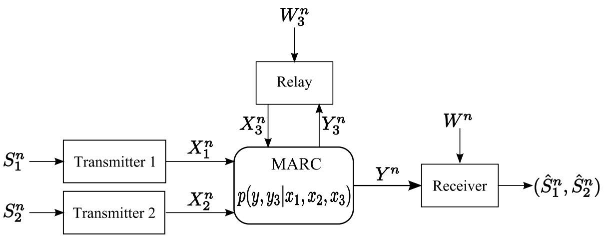

Fig. 1. The multiple-access-relay channel with correlated side information. $(\hat{S}_1^n,\hat{S}_2^n)$ are the reconstructions of (S_1^n,S_2^n) at the destination.

$p_X(x)$ to denote the probability mass function of a discrete RV X. We use boldface letters, e.g., $\mathbf{x}$, to denote vectors, and doublestroke font to denote matrices , e.g., $\mathbb{P}$. $H(\cdot)$ is used to denote the entropy of a discrete RV, $I(\cdot;\cdot)$ is used to denote the mutual information between two RVs, see [10, Ch. 2], and $X \leftrightarrow Y \leftrightarrow Z$ is used to denote a Markov relationship between X, Y and Z, as defined in [10, Notation]. Finally, $X \perp\!\!\!\perp Y$ is used to denote that X and Y are statistically independent, $\mathcal{N}^+$ is used to denote the set of positive integers, and ϕ is used to denote the empty set.

The MARC consists of two transmitters, a receiver (destination) and a relay. Transmitter $i, i=1,2,$ observes source sequence S_i^n. The objective of the receiver is to losslessly reconstruct the source sequences observed by the two transmitters, with the help of the relay. The relay and the receiver each observes its own side information, denoted by W_3^n and W^n, respectively, see Figure 1. The source and side information sequences, $\{S_{1,k}, S_{2,k}, W_k, W_{3,k}\} \in \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{W} \times \mathcal{W}_3, k = 1,2,\ldots,n,$ are arbitrarily correlated at each sample index k according to the joint distribution $p(s_1,s_2,w,w_3)$, and independent across different sample indices. All nodes know this joint distribution.

For transmission, a DM MARC with inputs $X_i \in \mathcal{X}_i, i = 1,2,3,$ and outputs $Y \in \mathcal{Y}, Y_3 \in \mathcal{Y}_3$, is available. The MARC is causal and memoryless in the sense of [1, Eqn. (1)]. A source-channel code for the MARC with correlated side information consists of two encoding functions at the transmitters: $f_i^{(n)} : \mathcal{S}_i^n \mapsto \mathcal{X}_i^n, i = 1,2,$ a decoding function at the destination, $g^{(n)} : \mathcal{Y}^n \times \mathcal{W}^n \mapsto \mathcal{S}_1^n \times \mathcal{S}_2^n$, and a set of n causal encoding functions at the relay, $x_{3,k} = f_{3,k}^{(n)}(y_{3,1}^{k-1}, w_{3,1}^n), k = 1,2,\ldots,n$. Observe that the code construction restricts the valid channel input distributions to obey the Markov chain

$$X_1 \leftrightarrow S_1^n \leftrightarrow S_2^n \leftrightarrow X_2. \tag{1}$$

Let $\hat{S}_i^n, i=1,2,$ denote the reconstruction of S_i^n at the receiver. The average probability of error of a source-channel code for the MARC is defined as $P_e^{(n)} \triangleq \Pr\left((\hat{S}_1^n,\hat{S}_2^n) \neq (S_1^n,S_2^n)\right)$. The sources S_1 and S_2 can be *reliably transmitted* over the MARC if there exists a sequence of source-channel codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. Next, we recall some results from [4].

III. Preliminaries

Let $X \in \mathcal{X}$, and $Y \in \mathcal{Y}$, be two discrete random variables with finite cardinalities. The joint probability distribution matrix $\mathbb{P}_{XY}$ is defined as $\mathbb{P}_{XY}(i,j) \triangleq \Pr\left(X{=}x_i, Y{=}y_j\right), \{x_i\}_{i=1}^{|\mathcal{X}|} \in \mathcal{X}, \{y_j\}_{j=1}^{|\mathcal{Y}|} \in \mathcal{Y}$. The marginal distribution matrix of an RV X is defined as the diagonal matrix $\mathbb{P}_X$ such that $\mathbb{P}_X(i,i) = \Pr\left(X = x_i\right), x_i \in \mathcal{X}$; $\mathbb{P}_X(i,j) = 0, \quad i \neq j$. This marginal distribution can also be represented in a vector form denoted by $\mathbf{p}_X$. The i'th element of $\mathbf{p}_X$ is $\mathbf{p}_X(i) \triangleq \Pr\left(X = x_i\right)$. The conditional joint probability distribution matrix $\mathbb{P}_{XY|z}$ is defined similarly.

Let $\tilde{\mathbb{P}}_{XY} \triangleq \mathbb{P}_X^{-\frac{1}{2}} \mathbb{P}_{XY} \mathbb{P}_Y^{-\frac{1}{2}}$ denote the spectral representation of the matrix $\mathbb{P}_{XY}$, and define the vector $\tilde{\mathbf{p}}_X$ as $\tilde{\mathbf{p}}_X = \mathbf{p}_X^{\frac{1}{2}}$, where $\mathbf{p}_X^{\frac{1}{2}}$ stands for an element-wise square root of $\mathbf{p}_X$. The conditional distributions $\tilde{\mathbb{P}}_{XY|z}$ and $\tilde{\mathbf{p}}_{X|y}$ are defined similarly.

Note that not every matrix $\tilde{\mathbb{P}}_{XY}$ can correspond to a given joint distribution matrix $\mathbb{P}_{XY}$. This is because a valid joint distribution matrix $\mathbb{P}_{XY}$ must have all its elements to be nonnegative and add up to 1. A necessary and sufficient condition for $\tilde{\mathbb{P}}_{XY}$ to correspond to a joint distribution matrix $\mathbb{P}_{XY}$ is given in the following theorem.

Theorem. ([4, Thm. 1]) Let $\mathbb{P}_X$ and $\mathbb{P}_Y$ be a pair of marginal distributions. A nonnegative matrix $\mathbb{P}_{XY}$ is a joint distribution matrix with marginal distributions $\mathbb{P}_X$ and $\mathbb{P}_Y$ if and only if the singular value decomposition (SVD) of the nonnegative matrix $\tilde{\mathbb{P}}_{XY}$ satisfies

$$\tilde{\mathbb{P}}_{XY} = \mathbb{M}\mathbb{D}\mathbb{N}^T = \mathbf{p}_X^{\frac{1}{2}} \left(\mathbf{p}_Y^{\frac{1}{2}}\right)^T + \sum_{i=2}^{l} \sigma_i \boldsymbol{\mu}_i \boldsymbol{\nu}_i^T, \tag{2}$$

where $\mathbb{M} \triangleq [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots \boldsymbol{\mu}_l]$ and $\mathbb{N} \triangleq [\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \ldots \boldsymbol{\nu}_l]$ are two matrices such that $\mathbb{M}^T\mathbb{M} = \mathbb{N}^T\mathbb{N} = \mathbb{I}$, $\mathbb{D} \triangleq \text{diag}[\sigma_1, \sigma_2, \ldots, \sigma_l]$[1], where $l = \min\{|\mathcal{X}|, |\mathcal{Y}|\}$; $\boldsymbol{\mu}_1 = \mathbf{p}_X^{\frac{1}{2}}, \boldsymbol{\nu}_1 = \mathbf{p}_Y^{\frac{1}{2}}$, and $\sigma_1 = 1 \geq \sigma_2 \geq \cdots \geq \sigma_l \geq 0$. That is, all the singular values of $\tilde{\mathbb{P}}_{XY}$ are non-negative and smaller than or equal to 1, the largest singular value of $\tilde{\mathbb{P}}_{XY}$ is 1, and its corresponding left and right singular vectors are $\mathbf{p}_X^{\frac{1}{2}}$ and $\mathbf{p}_Y^{\frac{1}{2}}$.

Next, we define the set of all possible conditional distributions $p(x_1,x_2|s_{1,1},s_{2,1})$ satisfying the Markov chain (1):

$$\mathcal{B}_{X_1X_2|S_1S_2} \triangleq \left\{ \begin{array}{l} p(x_1,x_2|s_{1,1},s_{2,1}) : \\ \exists n \in \mathcal{N}^+, p(x_1|s_1^n), p(x_2|s_2^n) \\ \text{s.t. } p(x_1,x_2|s_{1,1},s_{2,1}) = \\ \quad \dfrac{\displaystyle\sum_{s_{1,2}^n \in \mathcal{S}_1^{n-1}, s_{2,2}^n \in \mathcal{S}_2^{n-1}} p(x_1|s_1^n)p(x_2|s_2^n)p(s_1^n,s_2^n)}{p(s_{1,1},s_{2,1})} \end{array} \right\},$$

where $p(s_1^n,s_2^n) = \prod_{k=1}^n p(s_{1,k},s_{2,k})$. Note that as n can be arbitrarily large, optimization over the set of all conditional distributions $p(x_1|s_1^n)$ and $p(x_2|s_2^n)$ for all positive integers n is computationally intractable. Therefore, we are interested in identifying a larger set of probability distributions, whose characterization does not depend on n. Since this new set includes the original set of probability distributions, the conditions will still be necessary, and the computations become tractable.

Let $\sigma_i(\mathbb{A})$ denote the i-th singular value of the matrix $\mathbb{A}$. The following theorem characterizes constraints on $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2})$, $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1}})$, $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{2,1}})$ and

[1] We use diag[$\mathbf{a}$] to denote a rectangular matrix whose diagonal elements are the elements of the vector $\mathbf{a}$, and all off-diagonal elements are zero.

$\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1},s_{2,1}})$, and thereby gives a necessary condition for the n-letter Markov chain (1):[2]

Theorem. ([4, Thm. 4]) Let (S_1^n, S_2^n) be a pair of length-n independent and identically distributed (i.i.d.) sequences and let the random variables X_1 and X_2 satisfy the Markov chain (1). Let $S_{1,k}$ and $S_{2,j}$ be arbitrary elements of $\mathbf{S}_{1,1}^n$ and $\mathbf{S}_{2,1}^n$ respectively, that is, $k, j \in \{1, 2, \ldots, n\}$, then

$$\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,k},s_{2,j}}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1S_2}), \quad 2 \leq i. \tag{3}$$

Now, we define the set $\mathcal{B}'_{X_1X_2|S_1S_2}$ as follows:

$$\mathcal{B}'_{X_1X_2|S_1S_2} \triangleq \left\{\begin{array}{l} p_{X_1,X_2|S_1,S_2}(x_1, x_2|s_{1,1}, s_{2,1}): \\ \forall (s_{1,1}, s_{2,1}) \in \mathcal{S}_1 \times \mathcal{S}_2 \\ \sigma_i(\tilde{\mathbb{P}}_{X_1X_2}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) \quad i \geq 2 \\ \sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1}}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) \quad i \geq 2 \\ \sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{2,1}}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) \quad i \geq 2 \\ \sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1},s_{2,1}}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) \quad i \geq 2 \end{array}\right\}.$$

Note that the set $\mathcal{B}'_{X_1X_2|S_1S_2}$ is invariant to the symbol index, that is, $s_{1,1}$ and $s_{2,1}$ can be replaced by $s_{1,k}$ and $s_{2,k}$ for any $k \in \{2, 3, \ldots, n\}$. [4, Thm. 4] gives a necessary condition for the n-letter Markov chain (1), and therefore $\mathcal{B}_{X_1X_2|S_1S_2} \subseteq \mathcal{B}'_{X_1X_2|S_1S_2}$. Furthermore, the set $\mathcal{B}'_{X_1X_2|S_1S_2}$ is characterized by the singular values of the matrices $\tilde{\mathbb{P}}_{X_1X_2}, \tilde{\mathbb{P}}_{X_1X_2|s_{1,1}}, \tilde{\mathbb{P}}_{X_1X_2|s_{2,1}}$ and $\tilde{\mathbb{P}}_{X_1X_2|s_{1,1},s_{2,1}}$. Therefore, while the set $\mathcal{B}_{X_1X_2|S_1S_2}$ is defined in terms of infinitely many conditions, the set $\mathcal{B}'_{X_1X_2|S_1S_2}$ is defined with only a finite number of spectral conditions.

IV. Necessary Conditions

In this subsection we present the three new sets of necessary conditions for reliable transmission of correlated sources over DM MARCs.

A. A MAC Bound

The first new set of necessary conditions is a reminiscent of the so-called "MAC bound" for the relay channel, [10, Ch. 16], and takes into account (1).

Theorem 1. Any source pair (S_1, S_2) that can be transmitted reliably over the DM MARC with receiver side information W, as defined in Section II, must satisfy:

$$H(S_1|S_2, W) \leq I(X_1, X_3; Y|X_2, S_2, W, Q) \tag{4a}$$
$$H(S_2|S_1, W) \leq I(X_2, X_3; Y|X_1, S_1, W, Q) \tag{4b}$$
$$H(S_1, S_2|W) \leq I(X_1, X_2, X_3; Y|W, Q), \tag{4c}$$

for a joint distribution that factorizes as

$$p(q)p(s_1, s_2, w)p(x_1, x_2|s_1, s_2, q)\times$$
$$p(x_3|x_1, x_2, s_1, s_2, q)p(y|x_1, x_2, x_3), \tag{5}$$

with $|\mathcal{Q}| \leq 4$, and for every $q \in \mathcal{Q}$,

$$p(x_1, x_2|s_1, s_2, Q = q) \in \mathcal{B}_{X_1X_2|S_1S_2} \subseteq \mathcal{B}'_{X_1X_2|S_1S_2}. \tag{6}$$

Proof: The proof is given in Appendix A. ■

If the sources and the side information obey the Markov chain $S_1 \leftrightarrow W \leftrightarrow S_2$, then we have the following corollary:

Corollary 1. Let (S_1, S_2) be a source pair that satisfies the Markov relationship $S_1 \leftrightarrow W \leftrightarrow S_2$. If (S_1, S_2) can be transmitted reliably over the DM MARC with receiver side information W, as defined in Section II, then (S_1, S_2, W) must satisfy the constraints:

$$H(S_1|W) \leq I(X_1, X_3; Y|X_2, S_2, W, Q) \tag{7a}$$
$$H(S_2|W) \leq I(X_2, X_3; Y|X_1, S_1, W, Q) \tag{7b}$$
$$H(S_1|W) + H(S_2|W) \leq I(X_1, X_2, X_3; Y|W, Q), \tag{7c}$$

for a joint distribution that factorizes as

$$p(q)p(w)p(s_1|w)p(s_2|w)p(x_1, x_2|s_1, s_2, q)\times$$
$$p(x_3|x_1, x_2, s_1, s_2, q)p(y|x_1, x_2, x_3), \tag{8}$$

with $|\mathcal{Q}| \leq 4$, and for every $q \in \mathcal{Q}$,

$$p(x_1, x_2|s_1, s_2, Q = q) \in \mathcal{B}_{X_1X_2|S_1S_2} \subseteq \mathcal{B}'_{X_1X_2|S_1S_2}. \tag{9}$$

Remark 1. [4, Thm. 2] implies that if $S_1 \leftrightarrow W \leftrightarrow S_2$ then $\sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) \leq \sigma_2(\tilde{\mathbb{P}}_{S_1W})\sigma_2(\tilde{\mathbb{P}}_{WS_2})$. Therefore, characterizing the set $\mathcal{B}'_{X_1X_2|S_1S_2}$ based on the Markov relationship $S_1 \leftrightarrow W \leftrightarrow S_2$, that is replacing $\sigma_2(\tilde{\mathbb{P}}_{S_1S_2})$ with $\sigma_2(\tilde{\mathbb{P}}_{S_1W})\sigma_2(\tilde{\mathbb{P}}_{WS_2})$, gives a less restrictive set of sufficient conditions.

Remark 2. Corollary 1 can be specialized to [11, Thm. 5.2], which established the optimality of separation for MAC with correlated sources and the Markov chain $S_1 \leftrightarrow W \leftrightarrow S_2$. This can be done by setting $X_3 = \phi$ and replacing (Q, W) with $\bar{Q}$. Then, following arguments similar to the proof of the converse part of [11, Thm. 5.2], it follows that (8) becomes $p(s_1, s_2, \bar{q})p(x_1|s_1, \bar{q})p(x_2|s_2, \bar{q})p(y|x_1, x_2)$. Furthermore, in this case, the right hand sides (RHSs) of (7) are maximized by $p(x_i|s_i, \bar{q}) = p(x_i|\bar{q}), i = 1, 2$, which implies that $p(x_1, x_2|s_1, s_2, Q = q) \in \mathcal{B}'_{X_1X_2|S_1S_2}$.

B. A Broadcast Bound

The next two new sets of necessary conditions are a reminiscent of the so-called "broadcast bound" for the relay channel [10, Ch. 16].

Proposition 1. Any source pair (S_1, S_2) that can be transmitted reliably over the DM MARC with relay side information W_3 and receiver side information W, as defined in Section II, must satisfy the constraints:

$$H(S_1|S_2, W, W_3) \leq I(X_1; Y, Y_3|S_2, X_2, W, V) \tag{10a}$$
$$H(S_2|S_1, W, W_3) \leq I(X_2; Y, Y_3|S_1, X_1, W, V) \tag{10b}$$
$$H(S_1, S_2|W, W_3) \leq I(X_1, X_2; Y, Y_3|W, V), \tag{10c}$$

for a joint distribution that factorizes as

$$p(v, s_1, s_2, w, w_3)p(x_1, x_2|s_1, s_2, v)\times$$
$$p(x_3|v)p(y, y_3|x_1, x_2, x_3), \tag{11}$$

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

Proof: The proof is given in Appendix B. ■

Remark 3. [4, Thm. 4] requires (S_1^n, S_2^n) to be a pair of *i.i.d sequences* of length n. However, as V^n is *not* an *i.i.d sequence* then (S_1^n, S_2^n, V^n) is *not* a triplet of *i.i.d sequences*. Hence, it is not possible to use the approach of [4] to tighten Prop. 1. It is possible, however, to establish a different set of "broadcast-type" necessary conditions which benefits from the results of [4]. This is stated in Thm. 2.

Theorem 2. Any source pair (S_1, S_2) that can be transmitted reliably over the DM MARC with relay side information W_3 and receiver side information W, as defined in Section II, must satisfy the following constraints:

[2]Here we present a simplified version of [4, Thm. 4].

$$H(S_1|S_2,W,W_3) \le I(X_1;Y,Y_3|S_2,X_2,X_3,W,Q) \quad (12a)$$
$$H(S_2|S_1,W,W_3) \le I(X_2;Y,Y_3|S_1,X_1,X_3,W,Q) \quad (12b)$$
$$H(S_1,S_2|W,W_3) \le I(X_1,X_2;Y,Y_3|X_3,W,Q), \quad (12c)$$

for a joint distribution that factorizes as

$$p(q)p(s_1,s_2,w,w_3)p(x_1,x_2|s_1,s_2,q)\times p(x_3|x_1,x_2,w_3,q)p(y,y_3|x_1,x_2,x_3), \quad (13)$$

with $|\mathcal{Q}| \le 4$, and for every $q \in \mathcal{Q}$,

$$p(x_1,x_2|s_1,s_2,Q=q) \in \mathcal{B}_{X_1X_2|S_1S_2} \subseteq \mathcal{B}'_{X_1X_2|S_1S_2}. \quad (14)$$

Proof: The proof follows similar arguments to the proofs of Thm. 1 and Prop. 1, thus, it is omitted here. ∎

C. Discussion

Remark 4. The necessary conditions of Thm. 1 can be interpreted as representing decoding the relay assisted transmission (i.e., multiple-antenna transmitter) at the destination based on the destination's channel output and side information. On the other hand, the necessary conditions of Prop. 1 and of Thm. 2 can be interpreted as representing decoding at the destination using the channel outputs and side information of both the destination and the relay (i.e., multiple-antenna receiver).

Remark 5. The following proposition was established in [9]:

Proposition 2. ([9, Prop. 1]) Any source pair (S_1,S_2) that can be transmitted reliably over the DM MARC with receiver side information W, must satisfy the following constraints:

$$H(S_1|S_2,W) \le I(X_1,X_3;Y|X_2) \quad (15a)$$
$$H(S_2|S_1,W) \le I(X_2,X_3;Y|X_1) \quad (15b)$$
$$H(S_1,S_2|W) \le I(X_1,X_2,X_3;Y), \quad (15c)$$

for some input distribution $p(x_1,x_2,x_3)$.

Thm. 1 establishes necessary conditions which are at least as tight as Prop. 2. Since conditioning reduces entropy the RHSs of (4) are smaller than or equal to the RHSs of (15). Furthermore, the constraint (1) is not accounted for in Prop. 2, but is accounted for in Thm. 1. Therefore, as the LHSs of (4) and (15) are the same, Thm. 1 provides necessary conditions which are at least as tight as Prop. 2.

Remark 6. For independent sources, that is $p(s_1,s_2) = p(s_1)p(s_2)$, and $\mathcal{W} = \mathcal{W}_3 = \phi$, a combination of Thm. 1 and Thm. 2 specializes to the cut-set bound [12, Thm. 1]. In this case, the RHSs of (12) are identical to the first term in the RHS of [12, Eqns. (7)], while the RHSs of (4) are identical to the second term in the RHS of [12, Eqns. (7)]. Furthermore, (5) and (13) are the same. Note also that, for independent sources, $\sigma_2(\tilde{\mathbb{P}}_{S_1S_2}) = 0$, which implies that $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2})$, $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1}})$, $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{2,1}})$ and $\sigma_i(\tilde{\mathbb{P}}_{X_1X_2|s_{1,1},s_{2,1}})$ are all equal to zero. Therefore, X_1 and X_2 are independent and conditions (6) and (14) always hold. Finally, letting $R_1 \triangleq H(S_1), R_2 \triangleq H(S_2)$ implies that $H(S_1,S_2) = R_1 + R_2$, and therefore a combination of Thm. 1 and Thm. 2, for independent sources, results in [12, Eqns. (7)].

Remark 7. Consider specializing Prop. 1, Thm. 1 and Thm. 2 to the MAC. For Prop. 1 and Thm. 2 this can be done by setting $\mathcal{X}_3 = \mathcal{Y}_3 = \mathcal{W}_3 = \phi$, while for Thm. 1 this requires setting $\mathcal{X}_3 = \phi$. In this case the conditions in (4), (10) and (12) are identical. However, note that in (11) a general joint distribution $p(v,s_1,s_2,w)$ is considered, while in (5) and (13) $Q \perp\!\!\!\perp (S_1,S_2,W)$. Moreover, the required Markov chain of (1) is not accounted for by the chain of Prop. 1, contrary to Thm. 1 and Thm. 2. Therefore, we conclude that when specialized to the MAC scenario, Thm. 1 and Thm. 2 give the same bound which is tighter than Prop. 1.

Setting $\mathcal{X}_3 = \mathcal{Y}_3 = \mathcal{W}_3 = \phi$ as well as $\mathcal{W} = \phi$, specializes the model to the MAC with no side information at the receiver. For this model, both Thm. 1 and Thm. 2 specialize to [4, Thm. 7], which establishes necessary conditions for the MAC with correlated sources.

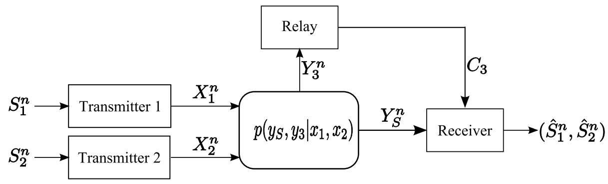

Fig. 2. Primitive semi-orthogonal MARC (PSOMARC).

V. A Numerical Example

We now demonstrate the improvement of Thm. 1 and Thm. 2 upon the cut-set bound [10, Ch. 18.1]. In order to simplify the arguments, we consider a scenario with no side information $\mathcal{W} = \mathcal{W}_3 = \phi$, and focus on the bound on $H(S_1,S_2)$. Before introducing the example we first recall the JSCC achievability scheme presented in [7, Thm. 1], for the case of $\mathcal{W} = \mathcal{W}_3 = \phi$. In particular, two of the six constraints in [7, Thm. 1] involve $H(S_1,S_2)$:

$$H(S_1,S_2) < I(X_1,X_2;Y_3|V_1,V_2,X_3) \quad (16a)$$
$$H(S_1,S_2) < I(X_1,X_2,X_3;Y), \quad (16b)$$

subject to a joint distribution that factorizes as

$$p(s_1,s_2)p(v_1)p(x_1|s_1,v_1)\times p(v_2)p(x_2|s_2,v_2)p(x_3|v_1,v_2)p(y_3,y|x_1,x_2,x_3). \quad (17)$$

Next, we recall the primitive semi-orthogonal MARC (PSOMARC) model depicted in Figure 2 [5]. This is a special MARC in which the relay-destination link is orthogonal to all the other channels. This link has a finite capacity, denoted by C_3. Even though the relay uses an orthogonal channel, this model still captures the main characteristics of the general MARC. We assume neither the relay nor the destination has side-information.

We consider a special PSOMARC and a source pair, for which we show that the cut-set bound fails to indicate whether reliable transmission is possible, while the new outer bounds we propose here do indicate that reliable transmission of the sources over the given channel is impossible.

Consider the PSOMARC defined by $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_3 = \mathcal{Y}_S = \{0,1\}$, and the channel transition probabilities detailed in Table I and II. We also let $C_3 = 0.1$. The cut-set bound constraint on the sum-rate of the PSOMARC, [5, Eqn. (9)], is given in (18) at the top of next page.

$Y_3 \setminus (X_1,X_2)$	(0,0)	(0,1)	(1,0)	(1,1)
0	0.87	0.25	0.51	0.24
1	0.13	0.75	0.49	0.76

TABLE I
THE TRANSITION PROBABILITY $p(y_3|x_1,x_2)$.

$$H(S_1,S_2) \le \mathrm{I}_{\text{cut-set}} \triangleq \max_{p(x_1,x_2)} \Big\{ I(X_1,X_2;Y_S) + \min\big\{C_3, I(X_1,X_2;Y_3|Y_S)\big\}\Big\}. \quad (18)$$

$$H(S_1,S_2) \le \mathrm{I}_{\text{new}} \triangleq \max_{p(x_1,x_2):\sigma_2(\tilde{\mathbb{P}}_{X_1X_2}) \le \sigma_2(\tilde{\mathbb{P}}_{S_1S_2})} \Big\{ I(X_1,X_2;Y_S) + \min\big\{C_3, I(X_1,X_2;Y_3|Y_S)\big\}\Big\}. \quad (19)$$

$$H(S_1,S_2) < \mathrm{I}_{\text{suff}} \triangleq \max_{X_1 \leftrightarrow S_1 \leftrightarrow S_2 \leftrightarrow X_2} \min\big\{ I(X_1,X_2;Y_3), I(X_1,X_2;Y_S) + C_3\big\}. \quad (20)$$

$Y \backslash (X_1,X_2)$	(0,0)	(0,1)	(1,0)	(1,1)
0	0.23	0.19	0.65	0.91
1	0.77	0.81	0.35	0.09

TABLE II
THE TRANSITION PROBABILITY $p(y|x_1,x_2)$.

Let (S_1,S_2) be a source-pair such that $\mathcal{S}_1 = \mathcal{S}_2 = \{0,1\}$, with the joint distribution $p(s_1,s_2)$ given in Table III.

$S_1 \backslash S_2$	0	1
0	0	0.04
1	0.045	0.915

TABLE III
THE JOINT DISTRIBUTION, $p(s_1,s_2)$ OF THE SOURCES.

For the transition probabilities defined in Tables I and II we have $\mathrm{I}_{\text{cut-set}} \approx 0.516$.[3] Next, consider relaxed versions of (4c) and (12c), with $\mathcal{W} = \mathcal{W}_3 = \phi$, specialized to the PSOMARC, given in (19) at the top of this page. Note that (19) is less restrictive than (4c) and (12c), as the maximization in (19) includes only the restriction due to $\tilde{\mathbb{P}}_{X_1X_2}$, while the restrictions due to the conditional distributions $\tilde{\mathbb{P}}_{X_1X_2|S_1}, \tilde{\mathbb{P}}_{X_1X_2|S_2}$ and $\tilde{\mathbb{P}}_{X_1X_2|S_1,S_2}$ are ignored. Finally, consider the sufficient conditions (16), specialized to the PSOMARC, given in (20) at the top of this page.

For the joint source distribution in Table III, we have $H(S_1,S_2) \approx 0.504$ and the sufficient condition (20) is evaluated as $\mathrm{I}_{\text{suff}} \approx 0.274$. Note that the cut-set necessary condition in (18) is larger than $H(S_1,S_2)$, and therefore it does not indicate whether these sources can be transmitted reliably. In contrast to (18), for the joint distribution given in Table III we have $\mathrm{I}_{\text{new}} \approx 0.485$. Hence, our new necessary condition in (19), explicitly indicates that reliable transmission is not possible.

This example demonstrates the improvement of Thm. 1 and Thm. 2 upon the cut-set bound.

Remark 8. This numerical example *does not follow immediately from the results of Kang and Ulukus for the MAC*, see [4, Subsection III.C]. To see this, consider the PSOMARC and sources defined in Tables I, II and III, and let $C_3 = 0.2$ (instead of 0.1). Here, (18) is evaluated as $\mathrm{I}_{\text{cut-set}} \approx 0.600$, while (19) is evaluated as $\mathrm{I}_{\text{new}} \approx 0.514$. Moreover, recall that $H(S_1,S_2) \approx 0.504$. Hence, for $C_3 = 0.2$, (19) fails to indicate whether reliable transmission of the sources is possible, while for $C_3 = 0.1$, (19) explicitly indicates that reliable transmission is impossible. This is in contrast to (18) which fails to indicate whether reliable transmission is possible, for both values of C_3.

[3]Throughout this section, the numerical values were found via exhaustive search. Note that the cut-set bound in (18) depends only on the channel transition probabilities and *not* on the joint distribution of the sources.

VI. CONCLUSIONS

We have derived three new sets of single-letter necessary conditions for reliable transmission of correlated sources over DM MARCs. We have shown that the new conditions are at least as tight as the previously known ones in the literature. One of the new sets is in the spirit of a MAC bound, while the other two sets follow from the broadcast bound. In two of the new sets of conditions we have exploited the Markov relationship between the sources and the channel inputs to restrict the feasible set of joint distributions, while still obtaining a computable characterization by constraining the possible input distributions in terms of their spectral properties rather than using n-letter constraints directly on the conditional distributions. Finally, we have demonstrated that the new necessary conditions improve upon the well known cut-set bound, constructing an explicit numerical example. Our results help in identifying the fundamental bounds on the level of cooperation that can be achieved in distributed source-channel communication networks.

APPENDIX A
PROOF OF THEOREM 1

Let $P_e^{(n)} \to 0$ for a sequence of encoders $f_i^{(n)}, i = 1,2,3$, and decoders $g^{(n)}$. We use Fano's inequality [10, Subsection 2.1], which states

$$\begin{aligned} H(S_1^n,S_2^n|\hat{S}_1^n,\hat{S}_2^n) &\le 1 + nP_e^{(n)}\log|\mathcal{S}_1 \times \mathcal{S}_2| \\ &\triangleq n\delta(P_e^{(n)}), \end{aligned} \quad \text{(A.1)}$$

where $\delta(x)$ is a non-negative function that approaches $\frac{1}{n}$ as $x \to 0$. We also obtain

$$\begin{aligned} H(S_1^n,S_2^n|\hat{S}_1^n,\hat{S}_2^n) &\overset{(a)}{\ge} H(S_1^n,S_2^n|\hat{S}_1^n,\hat{S}_2^n,W^n,Y^n) \\ &\overset{(b)}{=} H(S_1^n,S_2^n|W^n,Y^n) \\ &\overset{(c)}{\ge} H(S_1^n|S_2^n,W^n,Y^n), \end{aligned} \quad \text{(A.2)}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from the fact that $(\hat{S}_1^n,\hat{S}_2^n)$ is a deterministic function of (Y^n,W^n); and (c) follows from non-negativity of the entropy function for discrete sources. Constraint (4a) is a consequence of the following chain of inequalities:

$$\begin{aligned} &\sum_{k=1}^{n} I(X_{1,k},X_{3,k};Y_k|S_{2,k},X_{2,k},W_k) \\ &= \sum_{k=1}^{n} \Big[H(Y_k|S_{2,k},X_{2,k},W_k) \\ &\qquad - H(S_{2,k},Y_k|X_{1,k},X_{2,k},X_{3,k},W_k)\Big] \\ &\overset{(a)}{=} \sum_{k=1}^{n} \Big[H(Y_k|S_{2,k},X_{2,k},W_k) \\ &\qquad - H(Y_k|S_{2,k},X_{1,1}^k,X_{2,1}^k,X_{3,1}^k,W_k,Y^{k-1},Y_{3,1}^{k-1})\Big] \end{aligned}$$

$$\overset{(b)}{=}\sum_{k=1}^{n}\Big[H(Y_k|S_{2,k},X_{2,k},W_k) - H\big(Y_k|S_1^n,S_2^n,X_{1,1}^k,X_{2,1}^k,X_{3,1}^k, W^n,W_{3,1}^n,Y^{k-1},Y_{3,1}^{k-1}\big)\Big]$$
$$\overset{(c)}{\geq}\sum_{k=1}^{n}\Big[H(Y_k|S_2^n,X_{2,k},W^n,Y^{k-1}) - H(Y_k|S_1^n,S_2^n,W^n,W_{3,1}^n,Y^{k-1})\Big]$$
$$\overset{(d)}{=}\sum_{k=1}^{n}\Big[H(Y_k|S_2^n,W^n,Y^{k-1}) - H(Y_k|S_1^n,S_2^n,W^n,W_{3,1}^n,Y^{k-1})\Big]$$
$$=\sum_{k=1}^{n}I(S_1^n,W_{3,1}^n;Y_k|S_2^n,W^n,Y^{k-1})$$
$$\overset{(e)}{=}I(S_1^n,W_{3,1}^n;Y^n|S_2^n,W^n)$$
$$\overset{(f)}{\geq}I(S_1^n;Y^n|S_2^n,W^n)$$
$$=H(S_1^n|S_2^n,W^n)-H(S_1^n|S_2^n,W^n,Y^n)$$
$$\overset{(g)}{\geq}nH(S_1|S_2,W)-n\delta(P_e^{(n)}), \tag{A.3}$$
where (a) follows from the memoryless channel assumption; (b) follows from the causal Markov relationship $(S_1^n,S_2^n,W^n,W_{3,1}^n) \leftrightarrow (S_{2,k},X_{1,1}^k,X_{2,1}^k,X_{3,1}^k,W_k, Y^{k-1},Y_{3,1}^{k-1}) \leftrightarrow Y_k$ (see [13]); (c) follows from the fact that conditioning reduces entropy; (d) follows from the fact that $X_{2,k}$ is a deterministic function of S_2^n; (e) follows from the chain rule for mutual information; (f) follows from the chain rule for mutual information and the non-negativity of the mutual information; and (g) follows from the memoryless sources and side information assumption and from equations (A.1) and (A.2).

Following arguments similar to those that led to (A.3) we can also show
$$H(S_2|S_1,W) \leq \frac{1}{n}\sum_{k=1}^{n}I(X_{2,k},X_{3,k};Y_k|S_{1,k},X_{1,k},W_k)+\delta(P_e^{(n)})$$
$$H(S_1,S_2|W) \leq \frac{1}{n}\sum_{k=1}^{n}I(X_{1,k},X_{2,k},X_{3,k};Y_k|W_k)+\delta(P_e^{(n)}). \tag{A.4}$$
Note that the expressions: $I(X_{1,k},X_{3,k};Y_k|S_{2,k},X_{2,k},W_k)$, $I(X_{2,k},X_{3,k};Y_k|S_{1,k},X_{1,k},W_k)$, and $I(X_{1,k},X_{2,k},X_{3,k};Y_k|W_k)$, depend only on the marginal conditional distribution
$$p(x_{1,k},x_{2,k},x_{3,k}|s_{1,k},s_{2,k}) = p(x_{1,k},x_{2,k}|s_{1,k},s_{2,k})p(x_{3,k}|s_{1,k},s_{2,k},x_{1,k},x_{2,k})$$
with given $p(s_{1,k},s_{2,k},w_k)$ and $p(y_k|x_{1,k},x_{2,k},x_{2,k})$. Moreover, note that $X_{1,k}$ is a function of S_1^n while $X_{2,k}$ is a function of S_2^n. Hence, the Markov chains in (1) hold. Therefore it follows that
$$p(x_{1,k},x_{2,k}|s_{1,k},s_{2,k}) \in \mathcal{B}_{X_1X_2|S_1S_2} \subseteq \mathcal{B}'_{X_1X_2|S_1S_2}. \tag{A.5}$$

Next, we introduce the time-sharing random variable Q uniformly distributed over $\{1,2,\ldots,n\}$ and independent of all other random variables. We can write the following
$$\frac{1}{n}\sum_{k=1}^{n}I(X_{1,k},X_{3,k};Y_k|S_{2,k},X_{2,k},W_k) = \frac{1}{n}\sum_{k=1}^{n}I(X_{1,q},X_{3,q};Y_q|S_{2,q},X_{2,q},W_q,Q=k) = I(X_{1,Q},X_{3,Q};Y_Q|S_{2,Q},X_{2,Q},W_Q) = I(X_1,X_3;Y|S_2,X_2,W,Q), \tag{A.6}$$
where $X_1 \triangleq X_{1,Q}$, $X_2 \triangleq X_{2,Q}$, $X_3 \triangleq X_{3,Q}$, $Y \triangleq Y_Q$, $S_2 \triangleq S_{2,Q}$ and $W \triangleq W_Q$. Furthermore, since for all values of q we have $p(x_{1,q},x_{2,q}|s_{1,q},s_{2,q},Q=k)=p(x_{1,k},x_{2,k}|s_{1,k},s_{2,k})$ which satisfies (A.5), then we have that for $k=1,2,\ldots,n$ the following hold
$$p(x_{1,q},x_{2,q}|s_{1,q},s_{2,q},Q=k) \in \mathcal{B}'_{X_1X_2|S_1S_2}. \tag{A.7}$$
Finally, note that for all k, the expressions and structural constraints on the distribution chain are identical. Thus, from (A.3), (A.4) and (A.6) it follows that
$$H(S_1|S_2,W) \leq I(X_1,X_3;Y|S_2,X_2,W,Q)+\delta(P_e^{(n)})$$
$$H(S_2|S_1,W) \leq I(X_2,X_3;Y|S_1,X_1,W,Q)+\delta(P_e^{(n)})$$
$$H(S_1,S_2|W) \leq I(X_1,X_2,X_3;Y|W,Q)+\delta(P_e^{(n)}). \tag{A.8}$$
Finally, taking the limit as $n\rightarrow\infty$ leads to the constraints in (4).

Appendix B
Proof of Proposition 1

First, define the auxiliary RV:
$$V_k \triangleq (W_{3,1}^n, Y_{3,1}^{k-1}), \quad k=1,2,\ldots,n. \tag{B.1}$$
Now, constraint (10a) is a consequence of the following chain of inequalities:
$$\sum_{k=1}^{n}I(X_{1,k};Y_k,Y_{3,k}|S_{2,k},X_{2,k},W_k,V_k) \tag{B.2}$$
$$\overset{(a)}{=}\sum_{k=1}^{n}\Big[H(Y_k,Y_{3,k}|S_{2,k},X_{2,k},W_k,W_{3,1}^n,Y_{3,1}^{k-1}) - H(Y_k,Y_{3,k}|S_{2,k},X_{1,k},X_{2,k},W_k,W_{3,1}^n,Y_{3,1}^{k-1})\Big]$$
$$\overset{(b)}{=}\sum_{k=1}^{n}\Big[H(Y_k,Y_{3,k}|S_{2,k},X_{2,k},W_k,W_{3,1}^n,Y_{3,1}^{k-1}) - H(Y_k,Y_{3,k}|S_{2,k},X_{1,1}^k,X_{2,1}^k,X_{3,1}^k, W_k,W_{3,1}^n,Y^{k-1},Y_{3,1}^{k-1})]$$
$$\overset{(c)}{\geq}\sum_{k=1}^{n}\Big[H(Y_k,Y_{3,k}|S_2^n,X_{2,k},W^n,W_{3,1}^n,Y^{k-1},Y_{3,1}^{k-1}) - H(Y_k,Y_{3,k}|S_{2,k},X_{1,1}^k,X_{2,1}^k,X_{3,1}^k, W_k,W_{3,1}^n,Y^{k-1},Y_{3,1}^{k-1})\Big]$$

$$\overset{(d)}{=} \sum_{k=1}^{n} \Big[H(Y_k, Y_{3,k}|S_2^n, X_{2,k}, Y^{k-1}, W^n, W_{3,1}^n, Y_{3,1}^{k-1}) - H(Y_k, Y_{3,k}|S_1^n, S_2^n, X_{1,1}^k, X_{2,1}^k, X_{3,1}^k, W^n, W_{3,1}^n, Y^{k-1}, Y_{3,1}^{k-1})\Big]$$

$$\overset{(e)}{=} \sum_{k=1}^{n} \Big[H(Y_k, Y_{3,k}|S_2^n, W^n, W_{3,1}^n, Y^{k-1}, Y_{3,1}^{k-1}) - H(Y_k, Y_{3,k}|S_1^n, S_2^n, X_{1,1}^k, X_{2,1}^k, X_{3,1}^k, W^n, W_{3,1}^n, Y^{k-1}, Y_{3,1}^{k-1})\Big]$$

$$\overset{(f)}{\geq} \sum_{k=1}^{n} \Big[H(Y_k, Y_{3,k}|S_2^n, W^n, W_{3,1}^n, Y^{k-1}, Y_{3,1}^{k-1}) - H(Y_k, Y_{3,k}|S_1^n, S_2^n, W^n, W_{3,1}^n, Y^{k-1}, Y_{3,1}^{k-1})\Big]$$

$$= H(Y^n, Y_3^n|S_2^n, W^n, W_{3,1}^n) - H(Y^n, Y_3^n|S_1^n, S_2^n, W^n, W_{3,1}^n)$$

$$= I(S_1^n; Y^n, Y_3^n|S_2^n, W^n, W_{3,1}^n)$$

$$\overset{(g)}{\geq} I(S_1^n; Y^n|S_2^n, W^n, W_{3,1}^n)$$

$$= H(S_1^n|S_2^n, W^n, W_{3,1}^n) - H(S_1^n|S_2^n, W^n, W_{3,1}^n, Y^n)$$

$$\overset{(h)}{\geq} H(S_1^n|S_2^n, W^n, W_{3,1}^n) - H(S_1^n|S_2^n, W^n, Y^n)$$

$$\overset{(i)}{\geq} nH(S_1|S_2, W, W_3) - n\delta(P_e^{(n)}), \quad \text{(B.3)}$$

where (a) follows from (B.1); (b) follows from the fact that $X_{3,1}^k$ is a deterministic function of $(W_{3,1}^n, Y_{3,1}^{k-1})$ and from the memoryless channel assumption; (c) follows from the fact that conditioning reduces entropy; (d) follows from causality, [13]; (e) follows from the fact that $X_{2,k}$ is a deterministic function of S_2^n; (f) follows again from the fact that conditioning reduces entropy; (g) follows from the chain rule for mutual information and the nonnegativity of mutual information; (h) follows again from the fact that conditioning reduces entropy; and (i) follows from the memoryless sources and side information assumption and from equations (A.1) and (A.2).

Following arguments similar to those that led to (B.3) we can also show that

$$H(S_2|S_1, W, W_3) \leq \frac{1}{n}\sum_{k=1}^{n} I(X_{2,k}; Y_k, Y_{3,k}|S_{1,k}, X_{1,k}, W_k, V_k) + \delta(P_e^{(n)})$$

$$H(S_1, S_2|W, W_3) \leq \frac{1}{n}\sum_{k=1}^{n} I(X_{1,k}, X_{2,k}; Y_k, Y_{3,k}|W_k, V_k) + \delta(P_e^{(n)}). \quad \text{(B.4)}$$

Note that the mutual information expressions in (B.2) and (B.4) depend only on the marginal conditional distribution $p(x_{1,k}, x_{2,k}, x_{3,k}|s_{1,k}, s_{2,k}, w_k, v_k)$ with given $p(s_{1,k}, s_{2,k}, w_k, v_k)$ and $p(y_k, y_{3,k}|x_{1,k}, x_{2,k}, x_{3,k})$. Next we define the time-sharing random variable Q uniformly distributed over $\{1, 2, \dots, n\}$ and independent of all other random variables. We can write the following

$$\frac{1}{n}\sum_{k=1}^{n} I(X_{1,k}; Y_k, Y_{3,k}|S_{2,k}, X_{2,k}, W_k, V_k) = \frac{1}{n}\sum_{k=1}^{n} I(X_{1,q}; Y_q, Y_{3,q}|S_{2,q}, X_{2,q}, W_q, V_q, Q = k) = I(X_{1,Q}; Y_Q, Y_{3,Q}|S_{2,Q}, X_{2,Q}, W_Q, V_Q) = I(X_1; Y, Y_3|S_2, X_2, W, V), \quad \text{(B.5)}$$

where $X_1 \triangleq X_{1,Q}$, $X_2 \triangleq X_{2,Q}$, $Y \triangleq Y_Q$, $Y_3 \triangleq Y_{3,Q}$, $S_2 \triangleq S_{2,Q}$, $W \triangleq W_Q$ and $V \triangleq (V_Q, Q)$. Note that $(X_{1,k}, X_{2,k})$ and $X_{3,k}$ are independent given $(S_{1,k}, S_{2,k}, V_k)$. Moreover, the following Markov chain hold $(S_{1,k}, S_{2,k}) \leftrightarrow V_k \leftrightarrow X_{3,k}$. Therefore, for $\bar{v} = (v, k)$ we have

$$\Pr\{X_1 = x_1, X_2 = x_2, X_3 = x_3|S_1 = s_1, S_2 = s_2, V = \bar{v}\} = \Pr\{X_1 = x_1, X_2 = x_2|S_1 = s_1, S_2 = s_2, V = \bar{v}\} \times \Pr\{X_3 = x_3|V = \bar{v}\}. \quad \text{(B.6)}$$

Hence, the probability distribution is of the form given in (11). From (B.3)–(B.5), we have

$$H(S_1|S_2, W, W_3) \leq I(X_1; Y, Y_3|S_2, X_2, W, V) + \delta(P_e^{(n)})$$
$$H(S_2|S_1, W, W_3) \leq I(X_2; Y, Y_3|S_1, X_1, W, V) + \delta(P_e^{(n)})$$
$$H(S_1, S_2|W, W_3) \leq I(X_1, X_2; Y, Y_3|W, V) + \delta(P_e^{(n)}).$$

Finally, taking the limit as $n \to \infty$ and letting $P_e^{(n)} \to 0$ leads to the constraints in (10).

REFERENCES

[1] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks". *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[2] C. E. Shannon, "A mathematical theory of communication". *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and pp. 623–656, 1948.

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

[4] W. Kang and S. Ulukus, "A new data processing inequality and its applications in distributed source and channel coding". *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 56–69, Jan. 2011.

[5] R. Tandon and H. V. Poor, "On the capacity region of multiple-access relay channels". *Proc. Conf. Inform. Sci. and Syst.*, Baltimore, MD, Mar. 2011, pp. 1–5.

[6] Y. Murin, R. Dabora, and D. Gündüz, "Joint source-channel coding for the multiple-access relay channel". *Proc. IEEE Int. Symp. Inform. Theory*, Boston, MA, Jul. 2012, pp. 1947-1951.

[7] Y. Murin, R. Dabora, and D. Gündüz, "Mixed joint source-channel coding schemes for multiple-access-relay channels". *Proc. Int. Symp. Wireless Commun. Sys.*, Paris, France, Aug. 2012, pp. 994-998.

[8] H. E. Saffar, E. H. M. Alian, and P. Mitran, "Separation theorems for phase-incoherent multiple-user channels". Submitted to the *IEEE Trans. Inform. Theory*, Oct 2011, (arXiv:1110.3062).

[9] Y. Murin, R. Dabora, and D. Gündüz, "Source-channel coding for the multiple-access relay channel". *Proc. Int. Symp. Wireless Commun. Sys.*, Aachen, Germany, Nov. 2011, pp. 804-808.

[10] A. El. Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2012.

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

[12] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Hierarchical sensor networks: Capacity theorems and cooperative strategies using the multiple-access relay channel model". *Proc. 1st Conf. Sens. and Ad Hoc Commun. and Networks*, Santa Clara, CA, Oct. 2004, pp. 191–199.

[13] J. L. Massey, "Causality, feedback and directed information". *Proceedings of the IEEE International Symposium on Information Theory and its Applications*, Nov. 1990, Waikiki, HI, pp. 303-305.