

## Separating Distributed Source Coding From Network Coding

Aditya Ramamoorthy, *Member, IEEE*,  
Kamal Jain, Philip A. Chou, *Fellow, IEEE*, and  
Michelle Effros, *Senior Member, IEEE*

**Abstract**—This correspondence considers the problem of distributed source coding of multiple sources over a network with multiple receivers. Each receiver seeks to reconstruct all of the original sources. The work by Ho *et al.* 2004 demonstrates that random network coding can solve this problem at the potentially high cost of jointly decoding the source and the network code. Motivated by complexity considerations we consider the performance of separate source and network codes. Previous work by Effros *et al.* 2003 demonstrates the failure of separation between source and network codes for nonmulticast networks. We demonstrate that failure for multicast networks. We study networks with capacity constraints on edges. It is shown that the problem with two sources and two receivers is always separable. Counterexamples are presented for other cases.

**Index Terms**—Distributed source coding, multicast, network coding, separation.

### I. INTRODUCTION

The Slepian–Wolf (S–W) theorem [1] states that the lossless compression of two correlated sources that do not communicate with each other can be as efficient as the compression of the two sources that do communicate with each other. Csiszár shows in [2] that linear codes are sufficient to achieve the S–W bounds and computes error-exponents for various decoders. In that paper, he also shows the existence of a universal decoder that successfully decodes without requiring the knowledge of the joint statistics of the sources. In recent years there has been a flurry of activity (see [3]–[5] and their references) on code design for this distributed compression problem (hereafter referred to as the S–W problem), spurred mainly by applications in sensor networks and video coding problems.

The field of network coding investigates network flow problems when intermediate nodes in the network have the ability to forward functions of received packets rather than forwarding the packets unchanged. The seminal work of Ahlswede *et al.* [6] shows that network coding achieves the capacity of single-source, multiple-terminal multicast. Subsequent work [7], [8] shows that linear network coding is sufficient to achieve this capacity. Moreover, random linear network coding, which is an efficient distributed strategy, achieves this capacity with high probability [9], [10]. The multicast capacity of large random networks is considered in [11]. Variants of this problem involving multiple sources and multiple receivers are significantly harder and far less is known about them.

It is important to note that the classical S–W problem does not consider the sources to be communicating over a network, i.e., there is a direct edge from each source to the receiver. In addition the edges

do not have capacities on them. The S–W problem over a network is considered by Cristescu *et al.* [12] in the context of one receiver but they impose costs on edges rather than considering capacities on edges. In practical applications such as sensor networks, however, one would expect that the sources communicate over a network with capacities on the edges. This makes the problem of deciding the feasibility of a given distributed source coding problem with multiple sources and multiple receivers an interesting and important one. This problem is considered by Ho *et al.* [13]. They show by using the approach pioneered by Csiszár that as long as the minimum cuts between all nonempty subsets of sources and a particular receiver are larger than the corresponding conditional entropies (more details follow), random linear network coding followed by appropriate decoding at the receivers achieves (arbitrarily closely) the S–W bounds.

From a practical perspective one would like to leverage existing solutions to the classical S–W problem and, thus, separate the problem of sending the appropriate number of coded bits over a network from the source coding part. The solution proposed by [13] comes at the potentially high cost of jointly decoding the source and the network code. In general, the network code may destroy the structure in the source coder that allows tractable decoding. Indeed, if random network coding is used then this may happen with high probability.

This correspondence formally defines the problem of separation between distributed source coding and network coding and investigates the conditions under which separation holds.

Section II starts with a brief overview of distributed source coding and network coding and then provides a formal definition of separation between distributed source coding and network coding. Section III presents results on separation for networks with capacity constraints on edges and Section IV outlines the conclusions.

### II. OVERVIEW AND PROBLEM FORMULATION

Slepian and Wolf [1] in their landmark paper show that independent source coding of correlated sources (Fig. 1) can be as efficient as joint coding. For the case of two sources ( $X_1$  and  $X_2$ ), they show that if rates  $R_1$  (for  $X_1$ ) and  $R_2$  (for  $X_2$ ) satisfy

$$R_1 > H(X_1/X_2) \quad (1)$$

$$R_2 > H(X_2/X_1) \quad (2)$$

and

$$R_1 + R_2 > H(X_1, X_2) \quad (3)$$

then there exist independent source encoders at rates  $R_1$  (for  $X_1$ ) and  $R_2$  (for  $X_2$ ) and a decoding strategy that can recover  $X_1$  and  $X_2$  with high probability. Csiszár [2] shows that the S–W bounds can be achieved arbitrarily closely using random linear codes, by interpreting the source alphabets to be embedded in a Galois field  $\mathbb{F}_q$  of size  $q$ . Effros *et al.* [14] present a probabilistic statement of this fact if  $X_1$  and  $X_2$  are binary (i.e.,  $q = 2$ ). Specifically, let  $x_1^n$  and  $x_2^n$  represent vectors of elements in  $\mathbb{F}_2$  from sources  $X_1$  and  $X_2$  over  $n$  time units, and let  $A_{1,n}$  and  $A_{2,n}$  be  $[nR_1] \times n$  and  $[nR_2] \times n$  matrices, respectively, with elements in  $\mathbb{F}_2$ . Define source encoders  $\alpha_{1,n}(x_1^n) = A_{1,n}x_1^n$  and  $\alpha_{2,n}(x_2^n) = A_{2,n}x_2^n$ , and let  $\beta_n : \mathbb{F}_2^{[nR_1]} \times \mathbb{F}_2^{[nR_2]} \rightarrow \mathbb{F}_2^n \times \mathbb{F}_2^n$  denote the decoding function, which takes as input the encoded sources  $(\alpha_{1,n}(x_1^n), \alpha_{2,n}(x_2^n))$  and outputs an estimate of the original sources  $(x_1^n, x_2^n)$ . The decoding function  $\beta_n$  is a typical-set based decoder. (Details can be found in [14].) Let  $P_e(A_{1,n}, A_{2,n}) = Pr(\beta_n(\alpha_{1,n}(x_1^n), \alpha_{2,n}(x_2^n)) \neq (x_1^n, x_2^n))$  denote the error probability for the source code pair  $(A_{1,n}, A_{2,n})$ .

Manuscript received March 14, 2005; revised February 7, 2006. The material in this paper was presented in part at the 42nd Allerton Conference on Communication, Control and Computing, Monticello, IL, October 2004.

A. Ramamoorthy was with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA. He is now with Marvell Semiconductor Inc., Santa Clara, CA 95054 USA (e-mail: adityar@marvell.com).

K. Jain and P. A. Chou are with Microsoft Research, Redmond, WA 98052 USA (e-mail: kamalj@microsoft.com; pachou@microsoft.com).

M. Effros is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: effros@caltech.edu).

Communicated by R. W. Yeung, Guest Editor.

Digital Object Identifier 10.1109/TIT.2006.874534

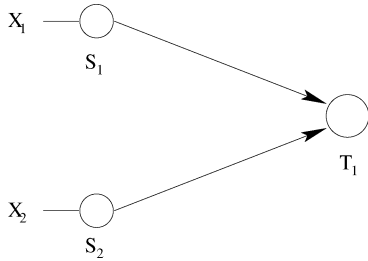


Fig. 1. Sources  $X_1$  and  $X_2$  being encoded independently at source encoders  $S_1$  and  $S_2$  and being sent to a terminal  $T_1$ .

**Theorem 1:** [14] Let  $(X_{1,1}, X_{2,1}), (X_{1,2}, X_{2,2}), \dots$  be drawn i.i.d. according to a joint distribution  $p(x_1, x_2)$  on  $\mathbb{F}_2 \times \mathbb{F}_2$ . Choose the sequence  $\{(A_{1,n}, A_{2,n})\}_{n=1}^{\infty}$  of rate- $(R_1, R_2)$  source encoders by choosing each entry of each matrix to be 0 or 1 with probability  $1/2$ . Then for an appropriately defined decoder and rates  $R_1$  and  $R_2$  that satisfy the S-W inequalities (1)–(3),  $E[P_e(A_{1,n}, A_{2,n})] \rightarrow 0$  as  $n \rightarrow \infty$ .

Intuitively, the result says that as long as the decoder receives a sufficient number of linearly combined bits from each source encoder it can correctly decode with high probability.

For nonbinary sources, Csiszár [2] embeds each source symbol  $X_i$  in a Galois field  $\mathbb{F}_q$  for some  $q \geq 2$ , and chooses the elements of the encoding matrices uniformly at random from  $\mathbb{F}_q$ . An alternative approach by Ho *et al.* [13], which aligns more closely with the network model used in this correspondence, is to embed each source symbol  $X_i$  in a vector space  $\mathbb{F}_2^{r_i}$  for some  $r_i \geq 1$ . That is,  $X_i$  produces a vector of  $r_i$  bits per unit time. Over  $n$  time units, this produces  $nr_i$  bits per block, which is then encoded using a random *binary* matrix of dimension  $\lceil nR_i \rceil \times nr_i$ . As  $n$  gets large, all that is required for error-free decoding with high probability is that the decoder receive the right number of bits generated by rows of these random matrices. Complex decoders such as maximum entropy or maximum *a posteriori* decoders are assumed in all three of these works.

To the best of our knowledge, Wyner was the first to propose a constructive approach for the S-W problem in [15]. It was used by Pradhan and Ramchandran [16] for practical S-W code design when the correlation between  $X$  and  $Y$  can be modeled by a binary symmetric channel, i.e., when  $Y$  can be considered to be a noisy version of  $X$  and the noise can be modeled as a BSC. The approach proceeds by encoding each source as a syndrome of an appropriately chosen binary code. The decoder then decodes the sources based on the values of the received syndromes.

Now suppose that the two sources are communicating with the decoder over a network with noiseless links. Fig. 2 depicts this situation. Sources  $X_1$  and  $X_2$  are observed at nodes  $S_1$  and  $S_2$  of the network. For now, suppose that only terminal node  $T_1$  is interested in reconstructing  $X_1$  and  $X_2$ . Without loss of generality one can assume that there exist virtual source encoders  $S'_1$  and  $S'_2$  that encode  $n$  source symbols, which are then fed to  $S_1$  and  $S_2$ . Suppose also that these virtual source encoders implement randomly chosen linear transformations as described earlier. Let the number of input bits into  $S_1$ , respectively,  $S_2$  be  $\lceil nH(X_1) \rceil$ , respectively,  $\lceil nH(X_2) \rceil$ . A natural strategy for the solution of this distributed source coding problem (now over a network) is the following.

- 1) Compute the minimum cuts between nodes  $S_1$  and  $T_1$ ,  $S_2$  and  $T_1$  and  $\{S_1, S_2\}$  and  $T_1$ . If  $R_1$  and  $R_2$  represent the rates that can be transmitted from  $S_1$  to  $T_1$  and  $S_2$  to  $T_1$ , the minimum cuts define a capacity region  $C_{T_1}$

$$R_1 \leq \text{min-cut}(S_1, T_1)$$

$$R_2 \leq \text{min-cut}(S_2, T_1)$$

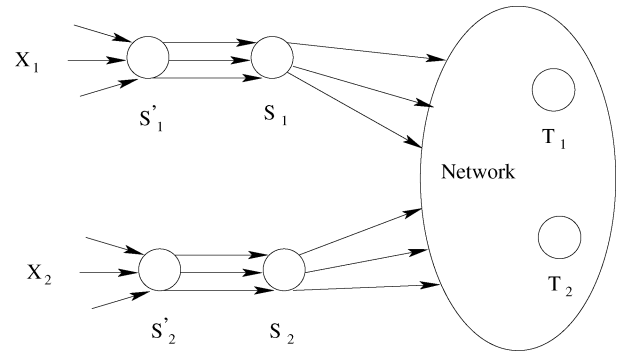


Fig. 2. Network with sources  $X_1$  and  $X_2$  being observed at source nodes  $S_1$  and  $S_2$  and two terminals  $T_1$  and  $T_2$ .  $S'_i$ 's ( $i = 1, 2$ ) can be thought of as virtual source encoders feeding coded bits to each source node. The network requires a transmission strategy that ensures that enough number of coded bits reach the terminals of interest.

and

$$R_1 + R_2 \leq \text{min-cut}(\{S_1, S_2\}, T_1). \quad (4)$$

If  $C_{T_1}$  has a nonempty intersection with the S-W region defined in (1) – (3), then the problem has a solution (for proofs of the necessity of this condition see [17] and [18]). We pick a rational rate vector  $(R_{S'_1}^{T_1}, R_{S'_2}^{T_1})$  that belongs to this intersection and choose  $n$  sufficiently large such that  $nR_{S'_1}^{T_1}$  and  $nR_{S'_2}^{T_1}$  are integers.

- 2) Over  $n$  time units, route the coded bits from  $S'_1$  and  $S'_2$  so that  $T_1$  receives  $nR_{S'_1}^{T_1}$  of the bits from  $S'_1$  and  $nR_{S'_2}^{T_1}$  of the bits from  $S'_2$ .

The decoder only needs to know the particular subsets of the bits from  $S'_1$  and  $S'_2$  that have been routed. It then just decodes as it would even if the sources were not operating over a network. By the proof of [14], [2] we can conclude that error-free decoding is possible with high probability at  $T_1$ . From a practical perspective this means that solutions such as those proposed by [16] that are based on linear codes would continue to work as long as a feasible rate vector  $(R_{S'_1}^{T_1}, R_{S'_2}^{T_1})$  exists and a suitable routing strategy is used.

Now consider what happens when we introduce another terminal  $T_2$ . That is, now the sources need to be decoded at two different terminals. We assume that the capacity region of  $T_2$  has a nonempty intersection with the S-W region of the sources; otherwise error-free decoding at  $T_2$  is not possible [17], [18]. Let us suppose that the terminals  $T_1$  and  $T_2$  can support feasible rate vectors  $(R_{S'_1}^{T_1}, R_{S'_2}^{T_1})$  and  $(R_{S'_1}^{T_2}, R_{S'_2}^{T_2})$ . There can be multiple strategies for the solution of this problem:

- 1) Transmit the coded bits via routing to both  $T_1$  and  $T_2$ .  
A routing strategy is required so that, as earlier,  $T_1$  receives a subset of size  $nR_{S'_1}^{T_1}$  bits from  $S'_1$  and a subset of size  $nR_{S'_2}^{T_1}$  bits from  $S'_2$ . A similar statement applies to  $T_2$  with the rate vector  $(R_{S'_1}^{T_2}, R_{S'_2}^{T_2})$ . The advantage of this solution would be that the decoders at  $T_1$  and  $T_2$  can be similar and will be directly decoding bits from  $S'_1$  and  $S'_2$ . In general, however, a routing strategy may not exist in some networks because of link sharing. For example, for the butterfly example of network coding in [6] no routing strategy can achieve a transmission rate of 2 bits per unit time.
- 2) Transmit the coded bits via random linear network coding to both  $T_1$  and  $T_2$ .

Ho *et al.* [13] show that if each node in the network performs random linear network coding then there exist decoding strategies at each terminal so that the sources can be decoded error-free with high probability. The problem with this approach is that the equivalent source code that needs to be decoded at a terminal is not under our control. Random network coding

allows combining of bits from the different source nodes and, thus, the original source-coded bits (from  $S'_1$  and  $S'_2$ ) may not be uniquely recoverable at the terminals. The equivalent source code would therefore lose any structure that allows tractable decoding. For example, if the original source codes were based on sparse parity-check matrix representations, e.g., LDPC codes, then the resultant parity-check matrices may end up becoming dense, ruling out the use of iterative decoding techniques. While this approach is promising in an information-theoretic sense, it fails to provide a practical solution to the problem. However, recent progress has been made in this direction [19].

- 3) Transmit the coded bits via appropriate network coding such that  $T_1$  and  $T_2$  can recover the original source-coded bits.

The approach here is to design a network code so that the recovery of a sufficient number of the original source-coded bits is possible at each terminal. The network code needs to support a connection such that terminal  $T_1$  recovers a subset of size  $nR_{S'_1}^{T_1}$  of the bits from  $S'_1$  and a subset of size  $nR_{S'_2}^{T_1}$  of the bits from  $S'_2$ . A similar statement should hold for  $T_2$ . If this can be achieved, as in the routing case, the decoders at the both the terminals will need only to decode subcodes of the original source codes. In a practical situation, the source code can be designed to allow tractable decoding, as well as good performance by utilizing advanced coding techniques such as LDPC codes. Thus, it is interesting to know whether network codes exist that are able to faithfully deliver the source coded bits as discussed earlier for the general case of multiple sources and receivers. This is the focus of this correspondence.

#### A. Formal Definition

The following discussion outlines the notation to be used in the rest of the correspondence and defines an instance of the distributed source coding problem over a network. We are given the following.

- a)  $N_S$  discrete memoryless sources denoted by  $X_i$ ,  $i = 1, \dots, N_S$ , whose output values are drawn i.i.d. from a joint distribution  $p(x_1, \dots, x_{N_S})$ . Each source alphabet is identified with a vector space over the binary field. This is without loss of generality since any discrete source alphabet can be considered to be a subset of a larger one by assigning zero probability to some symbols [2].
- b) A capacitated directed graph  $G = (V, E, C)$ , where  $V$  is a set of nodes,  $E$  is a set of directed edges, and  $C$  is a function that gives the capacity of each edge; all edge capacities are assumed to be rational. We are also given a set of source nodes  $S \subset V, |S| = N_S$ , and a set of receiver nodes  $T \subset V, |T| = N_R$ .

Given these items, we set up the problem as follows.

- a) We define the S–W region of the sources as

$$\mathcal{R}_{SW} = \{(R_1, R_2, \dots, R_{N_S}) : \forall B \subseteq \{1, 2, \dots, N_S\} \\ \sum_{i \in B} R_i > H(X_B/X_{B^c})\}$$

where  $X_B$  represents the vector of random variables  $(X_{i_1}, X_{i_2}, \dots, X_{i_{|B|}})$ , for  $i_k \in B, k = 1, \dots, |B|$ . We denote an instance of a distributed source coding problem over a network with capacity constraints by

$$P = \langle \mathcal{R}_{SW}, G, S, T \rangle.$$

- b) For each  $T_i \in T$  we define a capacity region with respect to  $S$ . This is the region that defines the maximum flow from each subset of  $S$  to the terminal  $T_i$

$$C_{T_i} = \{(R_1, R_2, \dots, R_{N_S}) : \forall B \subseteq S \\ \sum_{i \in B} R_i \leq \text{min-cut}(B, T_i)\}.$$

- c) We communicate  $n$  symbols in a block. This means that each source block  $X_i^n$  is encoded into  $\lceil nH(X_i) \rceil$  bits by its source encoder. Furthermore, edges with capacity  $C$  bits/symbol can transmit  $\lfloor nC \rfloor$  bits per block. Conceptually each edge can be regarded as multiple unit capacity edges, with each unit capacity edge capable of transmitting one bit per block. When communicating a block of length  $n$ , we consider the graph  $G^n = (V, E, \lfloor nC \rfloor)$ , or equivalently the graph  $(V, E_n, 1)$ , where  $E_n$  splits each edge from  $E$  into unit capacity edges.
- d) We introduce a set  $S'$  consisting of  $N_S$  virtual nodes denoted  $S'_1, \dots, S'_{N_S}$ , which can be regarded as source encoders, respectively, connected to  $S_1, \dots, S_{N_S}$ . Each encoder  $S'_i$  performs a binary linear encoding over a block of length  $n$  defined by a function  $\alpha_{i,n}$  that maps  $n$  consecutive samples of  $X_i$  (denoted by  $(X_{i,1}, X_{i,2}, \dots, X_{i,n})$ ) to the  $\lceil nH(X_i) \rceil$  bits (denoted by  $(U_{i,1}U_{i,2} \dots U_{i,\lceil nH(X_i) \rceil})$ ) used in their description. We define an augmented graph denoted by  $G'^n = (V \cup S', E_n \cup E'_n, 1)$ , where  $E'_n$  represents unit capacity edges from  $S'$  to  $S$  carrying the bits  $(U_{i,1}U_{i,2} \dots U_{i,\lceil nH(X_i) \rceil})$ ,  $i = 1, \dots, N_S$ .
- e) At any node in  $G'^n$ , the set of bits on edges entering the node is treated as a binary vector of dimension  $n_i$ , where  $n_i$  is the number of incoming edges at the node. Similarly the set of bits on edges leaving the node is treated as a binary vector of dimension  $n_o$ , where  $n_o$  is the number of outgoing edges at the node. The node implements a linear transformation specified by a binary matrix of dimension  $n_o \times n_i$ . Each row of this matrix corresponds to an outgoing edge  $e$ , and is said to be the local encoding vector for edge  $e$ . The set of all such linear transformations at the nodes of  $G'^n$ , or equivalently the set of all local encoding vectors on the edges of  $G'^n$ , represent a linear network code  $\alpha^n$  (having blocklength  $n$ ) for  $G'^n$ . A linear network code whose elements are chosen uniformly at random is a random linear network code.

*Definition 1: Feasibility.* Consider an instance of a distributed source coding problem over a network,  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$ . Let  $C_{T_i}$  be the capacity region of each receiver  $T_i \in T$  with respect to  $S$ . If

$$\mathcal{R}_{SW} \cap C_{T_i} \neq \emptyset, \forall i = 1, \dots, N_R \quad (5)$$

then the feasibility condition is said to be satisfied and  $P$  is said to be feasible.

*Theorem 2:* Sufficiency of the feasibility condition [13]. Consider an instance of a distributed source coding problem over a network,  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$ . If the feasibility condition (Definition 1) is satisfied, then a random linear network code for  $G'^n$  coupled with a minimum-entropy [2] or maximum-likelihood decoder at each receiver can recover the sources at each terminal in  $T$  with the probability of decoding error going to 0 as  $n \rightarrow \infty$ .

For the necessity of the feasibility condition we refer the reader to references [17] and [18].

#### B. Separation of Distributed Source Coding and Network Coding

In the sequel, we will work in the framework presented in Section II-B. As mentioned before, the result of Theorem 2 assumes the existence of a minimum-entropy/maximum-likelihood decoder that can be arbitrarily complex when random linear network codes are used. In this correspondence we study the feasibility of performing these operations separately. For this, we need a formal definition of separation between distributed source coding and network coding, which is presented.

*Definition 2: Separability.* Consider a distributed source coding problem over a network  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$ . Assume that

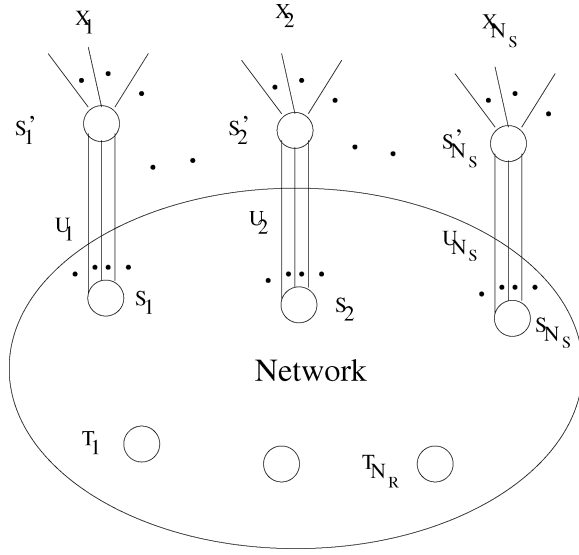


Fig. 3. Network with  $N_S$  sources ( $X_i$ 's), source encoders ( $S'_i$ 's) and source nodes ( $S_i$ 's). The source coded bits are represented by the  $U_i$ 's. There are  $N_R$  receivers ( $T_i$ 's).

$P$  is feasible. Hence there exist rate vectors  $(R_{S_1}^{T_i}, \dots, R_{S_{N_S}}^{T_i}) \in \mathcal{R}_{SW} \cap C_{T_i}$  for each  $T_i \in T$ . For a given blocklength  $n$ , let  $E_{S_j}^{T_i}$  denote a subset of the  $[nH(X_j)]$  edges in  $E'_n$  from  $S'_j$  to  $S_j$  with  $|E_{S_j}^{T_i}| \geq nR_{S_j}^{T_i}$ , for each  $S_j \in S$  and  $T_i \in T$ . Assume that all nodes are capable of linear network coding only.  $P$  is said to be *separable* if for any  $N_0$  there exist for some blocklength  $n \geq N_0$  a linear network code  $\alpha^n$ , rate vectors  $(R_{S_1}^{T_i}, \dots, R_{S_{N_S}}^{T_i})$  as earlier, and edge subsets  $E_{S_j}^{T_i}$  as earlier, such that the transfer function from the bits on  $E'_n$  to the bits on the input edges of  $T_i$  uniquely determines the bits on  $E_{S_1}^{T_i} \cup E_{S_2}^{T_i} \cup \dots \cup E_{S_{N_S}}^{T_i}$ , for each  $T_i \in T$ . We call such a linear network code  $\alpha^n$  *separable*.

Separability implies that we can achieve the optimal performance using separate source and network codes. The source encoder describes each source with a collection of bits. The network code losslessly delivers a sufficient subset of those bits to each receiver. The source decoder then recovers the source with high probability.

Note that even if  $P$  is separable, the terminals may be required to perform nonlinear operations in general for recovering the sources. However, if  $P$  is separable and practical distributed source codes for the classical S-W problem (based on linear codes) are used as the source encoders in the source nodes  $S'_1, \dots, S'_{N_S}$ , then the corresponding practical source decoders (outlined in [3] and [5]) can be used to recover the sources after a linear inversion of the network code.

### III. RESULTS FOR NETWORKS WITH CAPACITY CONSTRAINTS

In this section we present various results that characterize the separability of different distributed source coding problems over networks that have capacities on edges.

*Lemma 1:* Consider a problem  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$ , such that  $|T| = 1$ . If  $P$  is feasible, then  $P$  is separable.

*Proof:* Since  $P$  is feasible,  $\mathcal{R}_{SW} \cap C_{T_1} \neq \emptyset$ . Thus, for all sufficiently large  $n$ , in  $G'^n$  there exist a sufficient number of edge-disjoint paths from each source node to the terminal so that routing itself would suffice to ensure the delivery of information to the terminal at a rate vector that lies in the S-W region of the sources. Since in this case network coding is not needed, the received bits at the terminal  $T_1$  trivially determine the bits in  $E_{S_1}^{T_1} \cup E_{S_2}^{T_1} \dots E_{S_{N_S}}^{T_1}$  uniquely. Thus,  $P$  is separable.  $\square$

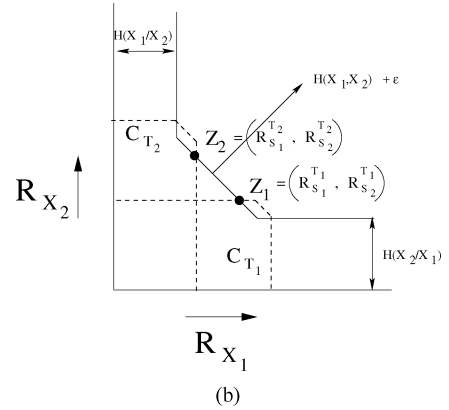
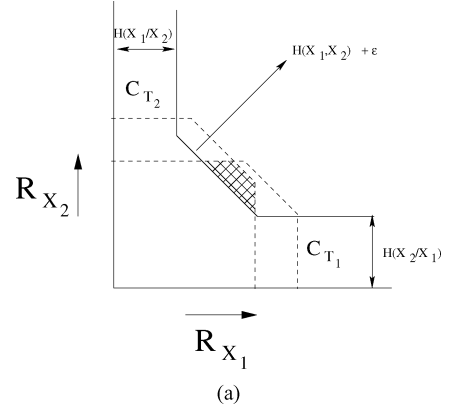


Fig. 4. The two regions defined by dotted lines are the capacity regions of  $T_1$  and  $T_2$ , respectively, and the region defined by solid lines is the S-W region of the sources. (a) The shaded region represents the region common to  $C_{T_1}$ ,  $C_{T_2}$  and  $\mathcal{R}_{SW}$ . (b) There is no point that is common to all three regions here.  $Z_2$  and  $Z_1$  are the closest operating points for each terminal on the S-W boundary.

The case corresponding to  $N_S = 1$  is not a distributed source coding problem since there is only one source. Nevertheless, we can see that this problem is separable in the sense of Definition 2 by the multicast result of Ahlswede *et al.* [6].

#### A. The Two-Sources, Two-Receivers Case

The following theorem shows that any feasible distributed source coding problem with two sources and two terminals is always separable.

*Theorem 3:* Consider a problem  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$ , with  $|S| = 2, |T| = 2$ . If  $P$  is feasible, then  $P$  is separable.

*Proof:* Since the connection is feasible we have  $\mathcal{R}_{SW} \cap C_{T_1} \neq \emptyset$  and  $\mathcal{R}_{SW} \cap C_{T_2} \neq \emptyset$ . There can be two cases as shown in Fig. 4(a) and (b).

a) Case 1)  $\mathcal{R}_{SW} \cap C_{T_1} \cap C_{T_2} \neq \emptyset$ .

This is the case shown in Fig. 4(a). Since  $\mathcal{R}_{SW}$  is an open set, there exists an open region in the intersection in which, for every sufficiently large  $n$ , there exists a single rate vector  $(R_{S_1}, R_{S_2})$  that can be supported at both terminals  $T_1$  and  $T_2$  in  $G'^n$ . Thus, the same set of bits can be sent to both  $T_1$  and  $T_2$  and the multicast result of [6] guarantees the existence of a  $P_{sol}^n$  such that  $P$  is separable.

b) Case 2)  $\mathcal{R}_{SW} \cap C_{T_1} \cap C_{T_2} = \emptyset$ .

The problem is more challenging when we consider the situation in Fig. 4(b). Unlike the earlier case, a single rate vector cannot be supported at both the terminals. Consequently the result of [6] no longer applies in a straightforward fashion. The proof that even this case is separable follows.

By the given conditions we can assume the existence of  $\epsilon > 0$  such that  $H(X_1, X_2) + \epsilon$  is rational and the line  $R_{X_1} + R_{X_2} = H(X_1, X_2) + \epsilon$  has a nonempty intersection with  $C_{T_1}$  and  $C_{T_2}$ . We force the terminals  $T_1$  and  $T_2$  to operate on the rational points marked  $Z_1 = (R_{S_1}^{T_1}, R_{S_2}^{T_1})$  and  $Z_2 = (R_{S_1}^{T_2}, R_{S_2}^{T_2})$ , respectively, in Fig. 4(b). Then, the following properties hold true.

- 1)
 
$$\begin{aligned} R_{S_1}^{T_1} &> R_{S_1}^{T_2} \\ R_{S_2}^{T_2} &> R_{S_2}^{T_1} \\ R_{S_1}^{T_i} + R_{S_2}^{T_i} &= H(X_1, X_2) + \epsilon, \quad \text{for } i = 1, 2. \end{aligned}$$
- 2) For  $Z'_1 \in C_{T_1} \cap \mathcal{R}_{SW} \cap \{(x_1, x_2) : x_1 + x_2 = H(X_1, X_2) + \epsilon\}$  and  $Z'_2 \in C_{T_2} \cap \mathcal{R}_{SW} \cap \{(x_1, x_2) : x_1 + x_2 = H(X_1, X_2) + \epsilon\}$

$$\text{dist}(Z_1, Z_2) \leq \text{dist}(Z'_1, Z'_2)$$

where  $\text{dist}(a, b)$  represents the distance between points  $a$  and  $b$ .

We choose  $n$  sufficiently large such that  $nR_{S_j}^{T_i}$  is integral for  $i, j = 1, 2$ . The proof is inspired by the technique used in [20].

For now let us only consider the paths from  $S_1$  to  $T_1$ , from  $S_1$  to  $T_2$ , and from  $S_2$  to  $T_2$ . For ease of explanation we let  $g = nR_{S_1}^{T_1}$ ,  $r_1 = nR_{S_1}^{T_2}$ , and  $r_2 = nR_{S_2}^{T_2}$ . Menger's theorem guarantees the existence of edge-disjoint paths in  $G'^n$  corresponding to these numbers. In particular, we denote by  $\mathbb{G}$  the set of  $g$  edge-disjoint paths from  $S_1$  to  $T_1$ , we denote by  $\mathbb{R}_1$  the set of  $r_1$  edge-disjoint paths from  $S_1$  to  $T_2$ , and we denote by  $\mathbb{R}_2$  the set of  $r_2$  edge-disjoint paths from  $S_2$  to  $T_2$ . Note that paths in  $\mathbb{R}_1 \cup \mathbb{R}_2$  are also edge-disjoint, and that  $g > r_1$ .

Each edge  $e$  in each path belonging to  $\mathbb{G} \cup \mathbb{R}_1 \cup \mathbb{R}_2$  is labeled (as explained later) by either one or two colors, namely *green* and/or *red*. Specifically, all edges in paths belonging to  $\mathbb{G}$  are labeled *green* and all edges in paths belonging to  $\mathbb{R}_1$  or  $\mathbb{R}_2$  are labeled *red*. Thus, edges in both  $\mathbb{G}$  and  $(\mathbb{R}_1 \cup \mathbb{R}_2)$  are labeled both *green* and *red*.

We claim that we can always find  $(g - r_1)$  exclusively *green* paths from  $S_1$  to  $T_1$ . To prove this, we define an algorithm  $A$  that takes as input a path  $P_1 \in \mathbb{G}$ .

#### Algorithm $A(P_1)$

- 1) Traverse  $P_1$  starting at node  $S_1$  and find the first edge  $e_1$  that is colored both *green* and *red*.
- 2) If no such  $e_1$  is found then **STOP**.
- 3) **ELSE** There are two possibilities.
  - a)  $e_1$  belongs to a path in  $\mathbb{R}_2$ .

We claim that this is impossible. To see this, suppose that  $e_1$  belonged to a path  $P' \in \mathbb{R}_2$  such that  $P' = P'_1 \rightarrow e_1 \rightarrow P'_2$ , where  $P'_1$  represents the portion of  $P'$  from  $S_2$  to  $e_1$  and  $P'_2$  represents the portion of  $P'$  from  $e_1$  to  $T_2$ .

We can color all edges on  $P_1$  from  $S_1$  to  $e_1$  *red* (in addition to their existing color *green*), and remove *red* from the color of edges in  $P'_1$ . That is, we can modify the solution to send one more bit of source  $S_1$  to receiver  $T_2$  through edge  $e_1$  and one fewer bit from source  $S_2$  to receiver  $T_2$ . This increases the rate from  $S_1$  to  $T_2$  by one bit per block and reduces the rate from  $S_2$  to  $T_2$  by one bit per block. The new rate vector  $(R_{S_1}^{T_2} + 1/n, R_{S_2}^{T_2} - 1/n)$  still lies on the line  $R_{X_1} + R_{X_2} = H(X_1, X_2) + \epsilon$ . But, this implies that  $Z_2$  and  $Z_1$  can be brought closer, which is a contradiction.

- b)  $e_1$  belongs to a path in  $\mathbb{R}_1$ .  
If  $e_1$  is the first edge of  $P_1$ , then **STOP**

**ELSE** Again suppose that  $e_1$  belonged to a path  $P' \in \mathbb{R}_1$ , such that  $P' = P'_1 \rightarrow e_1 \rightarrow P'_2$ , where  $P'_1$  represents the portion of  $P'$  from  $S_1$  to  $e_1$  and  $P'_2$  represents the portion of  $P'$  from  $e_1$  to  $T_2$ . Color all edges on  $P_1$  from  $S_1$  to  $e_1$  *red* (in addition to their existing color *green*), and remove *red* from the color of the edges in  $P'_1$ .

Now we define a condition that each path  $P_1 \in \mathbb{G}$  has to satisfy.

$$\begin{aligned} \text{Cond}(P_1) &= \{\text{All edges in } P_1 \text{ are } \textit{green}\} \\ &\text{or } \{\text{the first edge of } P_1 \text{ is } (\textit{green}, \textit{red})\}. \end{aligned} \quad (6)$$

We continue applying algorithm  $A$  to each path of  $\mathbb{G}$  until all paths in  $\mathbb{G}$  satisfy *Cond*. We claim that this process will eventually halt. To see this we define a function  $f(P_1)$  that given  $P_1 \in \mathbb{G}$  counts the number of (*green*, *red*) edges below the first set of contiguous (*green*, *red*) edges in  $P_1$ . Consider  $f_{\text{pot}} = \sum_{P \in \mathbb{G}} f(P)$ . Note that an application of  $A$  to a path from  $\mathbb{G}$  that violates *Cond* causes  $f_{\text{pot}}$  to strictly decrease. As a consequence, eventually all paths in  $\mathbb{G}$  will satisfy *Cond*.

At the end of this process, we claim that there exist  $(g - r_1)$  paths belonging to  $\mathbb{G}$  that are colored exclusively with *green*. This can easily be seen to be true, because if Algorithm  $A$  reroutes a path  $P' \in \mathbb{R}_1$ , then it removes the color *red* from one outgoing edge of  $S_1$  and places it on another outgoing edge. Thus, the number of outgoing edges that have the color *red* remains constant at  $r_1$ . Therefore, there have to be  $(g - r_1)$  outgoing edges that are purely *green*, which in turn means that there exist  $(g - r_1)$  paths from  $S_1$  to  $T_1$  that are exclusively *green*.

- a) Route  $n(R_{S_1}^{T_1} - R_{S_1}^{T_2})$  bits from  $S_1$  to  $T_1$ ;
- b) Similar argument shows that we can route  $n(R_{S_2}^{T_2} - R_{S_2}^{T_1})$  bits from  $S_2$  to  $T_2$ ;
- c) Each terminal needs exactly  $n(R_{S_1}^{T_2} + R_{S_2}^{T_1})$  bits more to satisfy its requirement. Invoking the multicast result of [6], we can send the same  $n(R_{S_2}^{T_2} - R_{S_2}^{T_1})$  bits to both receivers using network coding.

Thus, the two-sources, two-receivers problem can always be decomposed as depicted in Fig. 5 which in turn implies separability. The fact that a connection involving two sources and two receivers can be decomposed in two routed flows and one network coded flow was also noted in [21]. However, the proof presented is in the context of compressible sources and proceeds via a completely different approach that may be of independent interest.  $\square$

This result is somewhat surprising in light of the fact that we have source distributions and networks that serve as counterexamples for cases involving two sources and three receivers and three sources and two receivers. Those counterexamples follow.

#### B. The Two-Sources, Three-Receivers Case

Consider the network shown in Fig. 6(b) denoted by  $G$ . Here  $S = \{1, 2\}$  and  $T = \{9, 10, 11\}$ . Each edge has capacity  $(1 + \epsilon)$  bits. Fig. 6(a) shows the S-W region of the two sources ( $X_1$  and  $X_2$ ) denoted by  $\mathcal{R}_{SW} = \{(R_{X_1}, R_{X_2}) : R_{X_1} \geq 1, R_{X_2} \geq 1, R_{X_1} + R_{X_2} \geq 3\}$  and the capacity regions of the three terminals ( $T_1, T_2$ , and  $T_3$ ) given by

$$\begin{aligned} C_{T_1} &= \{(R_{X_1}, R_{X_2}) : R_{X_1} \leq 2 + 2\epsilon, R_{X_2} \leq 1 + \epsilon\} \\ C_{T_2} &= \{(R_{X_1}, R_{X_2}) : R_{X_1} \leq 1 + \epsilon, R_{X_2} \leq 2 + 2\epsilon\} \\ C_{T_3} &= \{(R_{X_1}, R_{X_2}) : R_{X_1} \leq 2 + 2\epsilon, R_{X_2} \leq 2 + 2\epsilon \\ &\quad R_{X_1} + R_{X_2} \leq 3 + 3\epsilon\}. \end{aligned} \quad (7)$$

Note that the S-W region includes its boundary in this problem. We make this choice for simplicity in what follows; it does not change the nature of the problem. We claim that the problem  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$  is not separable.

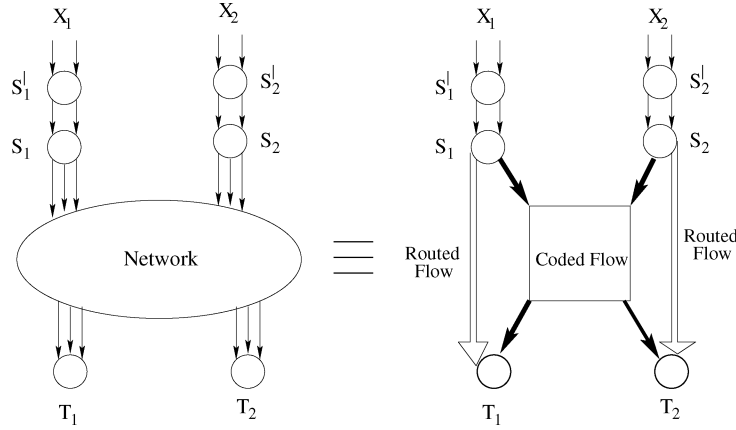


Fig. 5. Every two-source two-terminal distributed source coding problem over a network can be decomposed into one network coded flow (solid arrow) and two routed flows (unfilled arrows).

To prove this, we will assume that  $P$  is separable and derive a contradiction.

Assume that  $P$  is separable. By definition, there exists a block length  $n$  and a linear network code  $\alpha^n$  such that for each terminal  $T_i$ ,  $i = 1, 2, 3$ , there exist a rate vector  $(R_{S_1}^{T_i}, R_{S_2}^{T_i}) \in \mathcal{R}_{SW} \cap C_{T_i}$ , a subset  $E_{S_1}^{T_i}$  of the  $n[H(X_1)] = 2n$  edges from  $S_1'$  to  $S_1$ , and a subset  $E_{S_2}^{T_i}$  of the  $n[H(X_2)] = 2n$  edges from  $S_2'$  to  $S_2$  such that  $|E_{S_j}^{T_i}| \geq nR_{S_j}^{T_i}$  for  $j = 1, 2$  and the bits carried on the edges in these two subsets are uniquely determined by the bits on the edges entering  $T_i$ . In the following, we will let the linear network code  $\alpha^n$ , the rates  $\{R_{S_j}^{T_i}\}$ , and the edge sets  $\{E_{S_j}^{T_i}\}$  be any quantities satisfying these separability conditions.

Define the input vector into the network as  $U = [U_1^T U_2^T]^T$  such that  $U_1 = [U_{11}^T U_{12}^T]^T$  and  $U_2 = [U_{21}^T U_{22}^T]^T$  and  $\dim(U_{ij}) = n \times 1$  for  $i, j = 1, 2$ . For the linear network code  $\alpha^n$ , denote the net transformation from  $U$  to the input edges of each terminal  $T_i$  as  $Z_{T_i}$ . Further, denote the data vector flowing on each edge  $a \rightarrow b$  as  $X_{a-b}$ . For simplicity, we can assume the equalities  $X_{6-9} = X_{6-11} = X_{1-6}$ ,  $X_{3-9} = X_{1-3}$ ,  $X_{7-9} = X_{7-11} = X_{7-10} = X_{4-7}$ ,  $X_{5-10} = X_{2-5}$ , and  $X_{8-10} = X_{8-11} = X_{2-8}$ . To see that this is without loss of generality, let  $\alpha_1^n$  be an arbitrary linear network code satisfying the separability definition with rates  $\{R_{S_j}^{T_i}\}$  and edge sets  $\{E_{S_j}^{T_i}\}$ , such that

$$\begin{bmatrix} X_{6-9} \\ X_{3-9} \\ X_{7-9} \end{bmatrix} = \underbrace{\begin{bmatrix} H_{6-9} & 0 & 0 \\ 0 & H_{3-9} & 0 \\ 0 & 0 & H_{7-9} \end{bmatrix}}_H \begin{bmatrix} X_{1-6} \\ X_{1-3} \\ X_{4-7} \end{bmatrix} \quad (8)$$

where  $H_{a-b}$  denotes the  $n(1+\epsilon) \times n(1+\epsilon)$  transformation on edge  $a \rightarrow b$ . Consider the linear network code  $\alpha^n$  obtained from  $\alpha_1^n$  by replacing the matrix  $H$  by the identity matrix. This code then satisfies the separability definition with the same rates  $\{R_{S_j}^{T_i}\}$  and edges  $\{E_{S_j}^{T_i}\}$ , since for terminal  $T_1$  it is possible to uniquely determine the bits on the edges  $E_{S_1}^{T_1}$  and  $E_{S_2}^{T_1}$  from the bits on the edges entering  $T_1$  by applying the matrix  $H$  and then applying the decoder for the separable code  $\alpha_1^n$ . The other equality assumptions can be justified in a similar manner. Hence, if we show, by contradiction, that there cannot exist any separable linear network code satisfying the equality assumptions, then we have shown that there cannot exist any separable linear network code.

Now observe that the data vector  $X_{1-6}$  depends only on vector  $U_1$  and not on  $U_2$ . Therefore  $X_{1-6} = [A_{11} \ A_{12} \ 0 \ 0]U$ . Arguing in a similar fashion it is easy to see that

$$X_{1-3} = [A_{21} \ A_{22} \ 0 \ 0]U \quad (9)$$

$$X_{2-8} = [0 \ 0 \ B_{11} \ B_{12}]U \quad (10)$$

and

$$X_{2-5} = [0 \ 0 \ B_{21} \ B_{22}]U \quad (11)$$

where the  $A_{ij}$ 's and  $B_{ij}$ 's are the matrices that specify the transformation such that

$$\dim(A_{ij}) = \dim(B_{ij}) = n(1+\epsilon) \times n, \quad \forall i, j = 1, 2.$$

Continuing further note that

$$X_{3-4} = [H_{3-4}A_{21} \ H_{3-4}A_{22} \ 0 \ 0]U$$

and

$$X_{5-4} = [0 \ 0 \ H_{5-4}B_{21} \ H_{5-4}B_{22}]U.$$

It follows that

$$\begin{aligned} X_{4-7} &= [H_{4-7}^a H_{3-4}A_{21} \ H_{4-7}^a H_{3-4}A_{22} \\ &\quad H_{4-7}^b H_{5-4}B_{21} \ H_{4-7}^b H_{5-4}B_{22}]U \\ &= [M_{11} \ M_{12} \ M_{21} \ M_{22}]U \end{aligned} \quad (12)$$

where the  $\dim(M_{ij}) = n(1+\epsilon) \times n$  for all  $i, j = 1, 2$ .

Suppose that the matrices specify a solution  $P_{\text{sol}}^n$  that is separable. We will now consider each terminal and argue for the conditions that the matrices need to satisfy in order for separability to hold at them.

#### 1) Separability at terminal $T_1$

From the separability condition we know that there needs to exist a rate vector  $(R_{S_1}^{T_1}, R_{S_2}^{T_1}) \in \mathcal{R}_{SW} \cap C_{T_1}$ , a subset  $E_{S_1}^{T_1}$  of the  $n[H(X_1)] = 2n$  edges from  $S_1'$  to  $S_1$ , and a subset  $E_{S_2}^{T_1}$  of the  $n[H(X_2)] = 2n$  edges from  $S_2'$  to  $S_2$  such that  $|E_{S_j}^{T_1}| \geq nR_{S_j}^{T_1}$  for  $j = 1, 2$  and the bits carried on the edges in these two subsets are uniquely determined by the bits on the edges entering  $T_1$ .

From the capacity region  $C_{T_1}$  and the S-W region we know that  $1 \leq R_{S_2}^{T_1} \leq 1+\epsilon$  and  $1 \leq R_{S_1}^{T_1} \leq 2+2\epsilon$ . The S-W conditions also require that  $R_{S_1}^{T_1} + R_{S_2}^{T_1} \geq 3$ . Therefore we can conclude that for all rate vectors in  $\mathcal{R}_{SW} \cap C_{T_1}$  we have  $R_{S_2}^{T_1} \geq 1$  and  $R_{S_1}^{T_1} \geq 2-\epsilon$ . We will now demonstrate that these conditions translate into conditions on the ranks of certain submatrices.

The received vector at  $T_1$  is

$$[X_{6-9}^T \ X_{3-9}^T \ X_{7-9}^T]^T = [X_{1-6}^T \ X_{1-3}^T \ X_{4-7}^T]^T$$

(based on previous arguments). Therefore, the matrix  $Z_{T_1}$  can be written as

$$Z_{T_1} = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ M_{11} & M_{12} & M_{21} & M_{22} \end{bmatrix}. \quad (13)$$

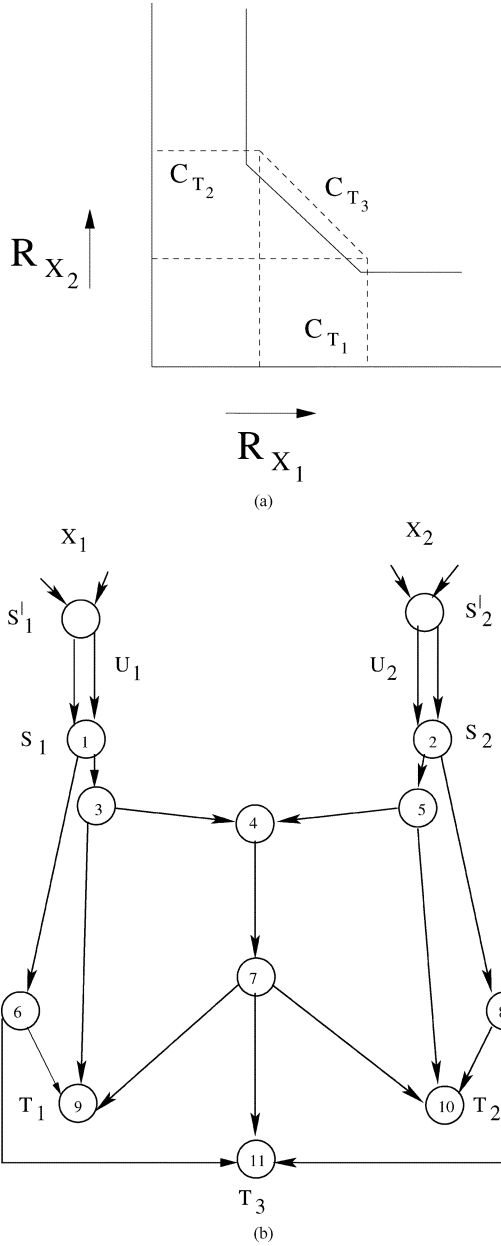


Fig. 6. (a) Capacity regions of the terminals (depicted by dotted lines) and the S-W region of the sources (depicted by solid lines). (b) Counterexample to separability for the case of two sources and three receivers.  $H(X_1) = H(X_2) = 2$ .  $H(X_1, X_2) = 3$ . The capacity of each edge equals  $1 + \epsilon$ .

At  $T_1$  we have access to the vector  $Z_{T_1}U$ . We note that by row operations the matrix can be transformed into a new matrix such that

$$Z_{T_1} \equiv \begin{bmatrix} A_{11} & A_{12} & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ 0 & 0 & M_{21} & M_{22} \end{bmatrix} \quad (14)$$

where the symbol  $\equiv$  denotes equivalence up to elementary row operations and premultiplication by a square nonsingular matrix since  $[M_{11} \ M_{12}] = H_{4-7}^a H_{3-4} [A_{21} \ A_{22}]$ . Therefore the submatrix of  $Z_{T_1}$  that specifies the transformation from  $U_1$  to  $T_1$  is given by

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \quad (15)$$

Since  $R_{S_1}^{T_1} \geq 2 - \epsilon$  this implies

$$\text{rank} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \geq (2 - \epsilon)n. \quad (16)$$

In the discussion that follows we will find an upper bound on the rank of  $M_{22}$ . The submatrix of  $Z_{T_1}$  that specifies the transformation from  $U_2$  to  $T_1$  is  $[M_{21} \ M_{22}]$  (from (14)). The submatrix  $M_{22}$  can be broken into two submatrices as  $M_{22} = [M_{22}^a \ M_{22}^b]$ .

Since  $R_{S_2}^{T_1} \geq 1$ ,  $|E_{S_2}^{T_1}| \geq n$ . We assume that the bits  $U_2$  have been suitably permuted so that the columns in  $[M_{21} \ M_{22}]$  correspond to the bits in the set  $E_{S_2}^{T_1}$  such that  $\text{rank}[M_{21} \ M_{22}] = |E_{S_2}^{T_1}|$ . We note that this also means that  $\text{rank}(M_{21}) = n$  since all columns in  $[M_{21} \ M_{22}]$  are linearly independent.

By the definition of separability the received bits at  $T_1$  need to uniquely determine the set of bits in  $E_{S_2}^{T_1}$ . Therefore, using Lemma 3 we have

$$\begin{aligned} \text{rank}[M_{21} \ M_{22}^a] + \text{rank}(M_{22}^b) \\ = \text{rank}[M_{21} \ M_{22}]. \end{aligned} \quad (17)$$

We know that  $\text{rank}[M_{21} \ M_{22}^a] = |E_{S_2}^{T_1}| \geq n$  and  $\text{rank}[M_{21} \ M_{22}] \leq n(1 + \epsilon)$  (the rank of a matrix is at most the number of rows). Therefore we conclude that  $\text{rank}(M_{22}^b) \leq n\epsilon$ .

Next we note that  $\text{rank}(M_{22}^a) \leq |E_{S_2}^{T_1}| - n$  (the rank of a matrix is at most the number of columns). Now  $|E_{S_2}^{T_1}| = nR_{S_2}^{T_1} \leq n(1 + \epsilon)$  (by the constraints of the capacity region  $C_{T_1}$ ), which implies  $\text{rank}(M_{22}^a) \leq n\epsilon$ . Finally we note that

$$\begin{aligned} \text{rank}(M_{22}) &= \text{rank}[M_{22}^a \ M_{22}^b] \\ &\leq \text{rank}(M_{22}^a) + \text{rank}(M_{22}^b) \\ &\leq 2n\epsilon. \end{aligned} \quad (18)$$

We will conclude this part of the proof by finding the rank of  $B_{21}$  and a lower bound on the rank of  $H_{4-7}^b H_{5-4}$ .

Note that  $M_{21} = H_{4-7}^b H_{5-4} B_{21}$ . Therefore

$$\begin{aligned} \text{rank}(M_{21}) \\ \leq \min(\text{rank}(H_{4-7}^b H_{5-4}), \text{rank}(B_{21})). \end{aligned} \quad (19)$$

Noting that  $\text{rank}(M_{21}) = n$  and that the column dimension of  $B_{21}$  is  $n$  we obtain

$$\text{rank}(B_{21}) = n \text{ and} \quad (20)$$

$$\text{rank}(H_{4-7}^b H_{5-4}) \geq n. \quad (21)$$

## 2) Separability at terminal $T_2$

Since the connectivity of terminals  $T_1$  and  $T_2$  is symmetric, by arguments analogous to the ones made previously we can obtain the following conclusions.

$$\text{rank} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \geq (2 - \epsilon)n \quad (22)$$

$$\text{rank}(M_{11}) \leq 2n\epsilon \quad (23)$$

$$\text{rank}(A_{22}) = n \quad (24)$$

$$\text{rank}(H_{4-7}^a H_{3-4}) \geq n. \quad (25)$$

For the sake of brevity we do not repeat the arguments.

## 3) Separability at terminal $T_3$

From the separability condition we know that there needs to exist a rate vector  $(R_{S_1}^{T_3}, R_{S_2}^{T_3}) \in \mathcal{R}_{SW} \cap C_{T_3}$  and a subset  $E_{S_1}^{T_3}$  of the  $n[H(X_1)] = 2n$  edges from  $S_1'$  to  $S_1$ , and a subset  $E_{S_2}^{T_3}$  of the  $n[H(X_2)] = 2n$  edges from  $S_2'$  to  $S_2$  such that  $|E_{S_j}^{T_3}| \geq nR_{S_j}^{T_3}$

for  $j = 1, 2$  and the bits carried on the edges in these two subsets are uniquely determined by the bits on the edges entering  $T_3$ . The region  $C_{T_3} \cap R_{SW}$  is defined by

$$\begin{aligned} C_{T_3} \cap R_{SW} = \{ & (R_{X_1}, R_{X_2}) : 1 \leq R_{X_1} \leq 2 + 2\epsilon, \\ & 1 \leq R_{X_2} \leq 2 + 2\epsilon, 3 \leq R_{X_1} + R_{X_2} \leq 3 + 3\epsilon \}. \end{aligned} \quad (26)$$

In this part of the proof we will use the previously derived conditions on the ranks of different submatrices and the conditions to arrive at a contradiction that will conclude the proof.

We will need two simple facts (proved in Lemma 2 in the Appendix) regarding the ranks of sums of matrices. If  $A$  and  $B$  are two matrices of the same dimension then

$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B) \quad \text{and} \quad (27)$$

$$\text{rank}(A + B) \geq \text{rank}(A) - \text{rank}(B). \quad (28)$$

We will first present an upper bound on the rank of the matrix  $B_{22}$ . Toward this end, we note that  $\text{rank}(H_{4-7}^b H_{5-4}) \geq n$  (from (21)). In the discussion that follows, we will assume that  $\text{rank}(H_{4-7}^b H_{5-4}) = n$ . It will be clear that this yields the weakest upper bound. We know that

$$M_{22} = H_{4-7}^b H_{5-4} B_{22}. \quad (29)$$

By elementary row operations we can put the right-hand side (RHS) of the equation in the form

$$\begin{bmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_{22}^1 \\ B_{22}^2 \end{bmatrix} \quad (30)$$

where the dimensions of  $Q_{11}$  are  $n \times n$  and the dimensions of  $Q_{12}$  are  $n \times n\epsilon$ . Further,  $\text{rank}[Q_{11} \ Q_{12}] = n$ . Without loss of generality, we assume that all of the columns of  $Q_{11}$  are linearly independent. If this is not the case, then we can first redefine  $Q_{11}$  to be the set of (linearly independent) columns and then redefine the block matrices  $B_{22}^1$  and  $B_{22}^2$  appropriately. Now

$$\begin{aligned} \text{rank}(M_{22}) &= \text{rank}(Q_{11} B_{22}^1 + Q_{12} B_{22}^2) \\ &\geq \text{rank}(Q_{11} B_{22}^1) - \text{rank}(Q_{12} B_{22}^2) \\ &\geq \text{rank}(Q_{11} B_{22}^1) - n\epsilon \\ &\quad (\text{since the number of rows in } B_{22}^2 \text{ is } n\epsilon) \\ &= \text{rank}(B_{22}^1) - n\epsilon \\ &\quad (\text{since } Q_{11} \text{ is an invertible matrix}). \end{aligned} \quad (31)$$

Noting that  $\text{rank}(M_{22}) \leq 2n\epsilon$  [from (18)], we obtain  $\text{rank}(B_{22}^1) \leq 3n\epsilon$ . Since  $\text{rank}(B_{22}^2) \leq n\epsilon$ , we finally obtain  $\text{rank}(B_{22}) \leq 4n\epsilon$ .

Our next step will be to compute a lower bound on the rank of  $B_{12}$ . We observe that

$$\begin{aligned} \text{rank} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\ \leq \text{rank} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} + \text{rank} \begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix} \\ \leq n + \text{rank}(B_{12}) + 4n\epsilon \\ \quad (\text{since } \text{rank}(B_{21}) = n, \text{rank}(B_{22}) \leq 4n\epsilon). \end{aligned} \quad (32)$$

Noting that

$$\text{rank} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \geq (2 - \epsilon)n$$

we conclude that

$$\text{rank}(B_{12}) \geq n(1 - 5\epsilon). \quad (33)$$

We will now determine an upper bound on  $|E_{S_1}^{T_3}| + |E_{S_2}^{T_3}|$  that will give us the required contradiction.

It is easy to see that

$$Z_{T_3} = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 \\ M_{11} & M_{12} & M_{21} & M_{22} \\ 0 & 0 & B_{11} & B_{12} \end{bmatrix}. \quad (34)$$

Notice that the rows corresponding to  $X_{7-11}$  have been moved to the middle of the matrix for notational convenience. Suppose that the first  $2n$  columns of  $Z_{T_3}$  are permuted so that the first  $|E_{S_1}^{T_3}|$  columns correspond to the bits in the set  $E_{S_1}^{T_3}$ . We write the new matrix  $Z'_{T_3}$  as

$$Z'_{T_3} = [H_1 \ H_2 \ H_3] \quad (35)$$

where  $\dim(H_1) = 3n(1 + \epsilon) \times |E_{S_1}^{T_3}|$ ,  $\dim(H_2) = 3n(1 + \epsilon) \times (2n - |E_{S_1}^{T_3}|)$ , and  $\dim(H_3) = 3n(1 + \epsilon) \times 2n$ . We make the following observations. Since the columns of  $H_1$  correspond to the bits in  $E_{S_1}^{T_3}$ , we need  $\text{rank}(H_1) = |E_{S_1}^{T_3}|$ . By the definition of separability the received bits at  $T_3$  need to uniquely determine the bits in  $E_{S_1}^{T_3}$ . Therefore, using Lemma 3, we can conclude that

$$\text{rank}(H_1) + \text{rank}[H_2 \ H_3] = \text{rank}(Z'_{T_3}) \quad (36)$$

where

$$H_3 = \begin{bmatrix} 0 & 0 \\ M_{21} & M_{22} \\ B_{11} & B_{12} \end{bmatrix}. \quad (37)$$

In the discussion later, we will present a lower bound on the rank of  $H_3$ . By elementary row operations that can be represented by premultiplication by a matrix of the form

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & R & I \end{bmatrix} \quad (38)$$

we can convert  $H_3$  into a matrix of the form

$$H'_3 = \begin{bmatrix} 0 & 0 \\ M_{21} & M_{22} \\ 0 & B'_{12} \end{bmatrix} \quad (39)$$

where  $R$  is chosen so that  $RM_{21} + B_{11} = 0$  (such an  $R$  exists since  $M_{21}$  has rank  $n$ ) and  $RM_{22} + B_{12} = B'_{12}$ . Now

$$\begin{aligned} \text{rank}(B'_{12}) &\geq \text{rank}(B_{12}) - \text{rank}(RM_{22}) \\ &\geq n(1 - 5\epsilon) - 2n\epsilon \text{ from } ((33) \text{ and } (18)) \\ &= n(1 - 7\epsilon). \end{aligned} \quad (40)$$

Consider the set of linearly independent columns in  $B'_{12}$ . We note that the corresponding columns in the matrix  $[M_{22}^T \ B'_{12}{}^T]^T$  are linearly independent of the columns in  $[M_{21}^T \ 0]^T$ . This is because a linear combination of columns in  $[M_{21}^T \ 0]^T$  necessarily has zeros in the last  $n(1 + \epsilon)$  coordinates.

Thus, we conclude that the total number of linearly independent columns in  $H_3$  or the rank of  $H_3$  is at least



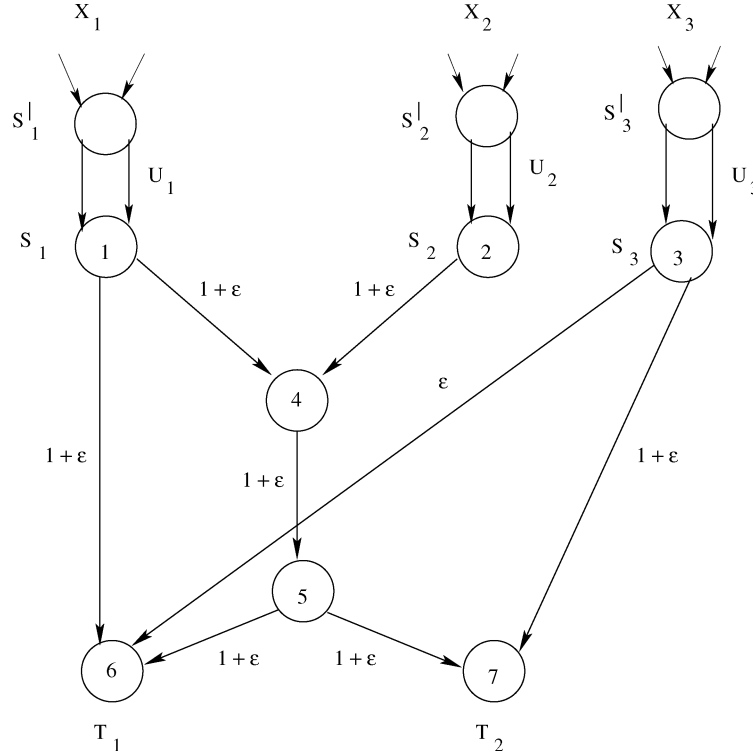


Fig. 7. Counterexample to separability for the case of three sources and two receivers. The capacity of edge  $3 \rightarrow 6$  is  $\epsilon$ . All other edges have capacity  $1 + \epsilon$ . The correlation model is explained in the text.

$$\text{rank}(M_{21}) + \text{rank}(B'_{12}) \geq n(2 - 7\epsilon).$$

Substituting in (36) we see that

$$\begin{aligned} \text{rank}(H_1) &= |E_{S_1}^{T_3}| = \text{rank}(Z_{T_3}^T) - \text{rank}[H_2 \ H_3] \\ &\leq 3n(1 + \epsilon) - n(2 - 7\epsilon) \\ &= n(1 + 10\epsilon). \end{aligned} \quad (41)$$

A similar argument shows that  $|E_{S_2}^{T_3}| \leq n(1 + 10\epsilon)$  giving us  $|E_{S_1}^{T_3}| + |E_{S_2}^{T_3}| \leq n(2 + 20\epsilon)$ . Therefore, if  $\epsilon < 1/20$  then the total number of bits that can be received at  $T_3$  is strictly less than  $3n$ . This provides the required contradiction and concludes our proof.  $\square$

### C. The Three-Sources, Two-Receiver Case

We can also find counterexamples to separability for the case of three sources and two terminals. Fig. 7 shows a network, denoted by  $G$ , along with three sources  $X_1$ ,  $X_2$ , and  $X_3$ . The source entropies are  $H(X_1) = 1$ ,  $H(X_2) = 1 + \epsilon_1$ , and  $H(X_3) = 1 + \epsilon_1$ , and the S-W region is

$$\mathcal{R}_{SW} = \{(R_{X_1}, R_{X_2}, R_{X_3}) : \quad (42)$$

$$R_{X_1} \geq H(X_1/X_2, X_3) = 1 \quad (43)$$

$$R_{X_2} \geq H(X_2/X_3, X_1) = \epsilon_1 \quad (44)$$

$$R_{X_3} \geq H(X_3/X_1, X_2) = \epsilon_1 \quad (45)$$

$$R_{X_1} + R_{X_2} \geq H(X_1, X_2/X_3) = 1 + \epsilon_1 \quad (46)$$

$$R_{X_2} + R_{X_3} \geq H(X_2, X_3/X_1) = 1 + 2\epsilon_1 \quad (47)$$

$$R_{X_3} + R_{X_1} \geq H(X_3, X_1/X_2) = 1 + \epsilon_1 \quad (48)$$

$$R_{X_1} + R_{X_2} + R_{X_3} \geq H(X_1, X_2, X_3) = 2 + 2\epsilon_1. \quad (49)$$

Here, we assume, as previously, that the boundary of the S-W region can be achieved. All edges in  $G$  have capacity  $1 + \epsilon$ , except edge  $3 \rightarrow 6$ , which has capacity  $\epsilon$ , where  $2\epsilon_1 > \epsilon > \epsilon_1 > 0$ . Here  $S = \{1, 2, 3\}$ ,  $T = \{6, 7\}$ . We claim that  $P = \langle \mathcal{R}_{SW}, G, S, T \rangle$  is not separable.

Again, to prove this, we will assume that  $P$  is separable, with a separable linear network code  $\alpha^n$ , and we will derive a contradiction.

Let the input vector into the network be denoted  $U = [U_1^T \ U_2^T \ U_3^T]^T$ , where  $\dim(U_1) = n \times 1$  and  $\dim(U_2) = \dim(U_3) = n(1 + \epsilon_1) \times 1$ . Without loss of generality, we can assume that in  $\alpha^n$ ,  $X_{5-6} = X_{5-7} = X_{4-5}$ . To see this, suppose that there exists a separable linear network code  $\alpha_1^n$  in which  $X_{5-6} = H_{5-6}X_{4-5}$ , where  $H_{5-6}$  is an arbitrary  $n(1 + \epsilon) \times n(1 + \epsilon)$  matrix. Then there is an equivalent separable linear network code in which  $X_{5-6} = X_{4-5}$ , for which the decoder at  $T_1$  must first multiply the vector received over edge  $5 \rightarrow 6$  by  $H_{5-6}$ . A similar argument shows that  $X_{5-7} = X_{4-5}$ .

Then

$$X_{1-6} = [A_{n(1+\epsilon) \times n} \ 0 \ 0]U$$

$$X_{3-6} = [0 \ 0 \ C_{n \times n(1+\epsilon_1)}]U,$$

$$X_{4-5} = [(B_1)_{n(1+\epsilon) \times n} \ (B_2)_{n(1+\epsilon) \times n(1+\epsilon_1)} \ 0]U$$

$$\text{and } X_{3-7} = [0 \ 0 \ D_{n(1+\epsilon) \times n(1+\epsilon_1)}]U$$

where the submatrices  $A$ ,  $B_1$ ,  $B_2$ ,  $C$ , and  $D$  specify the transformations, and their dimensions are specified by the appropriate subscripts. We will let  $Z_{T_i}$  specify the net transformation from  $U$  to the input edges of  $T_i$ .

Suppose that the matrices specify a solution  $P_{\text{sol}}^n$  that is separable. We will arrive at a contradiction for an appropriate range of  $\epsilon$ . The matrix  $Z_{T_1}$  can be written as

$$Z_{T_1} = \begin{bmatrix} A & 0 & 0 \\ B_1 & B_2 & 0 \\ 0 & 0 & C \end{bmatrix}. \quad (50)$$

By the constraints of  $\mathcal{R}_{SW}$ , we need

$$\begin{aligned} R_{S_2}^{T_1} + R_{S_3}^{T_1} &\geq 1 + 2\epsilon_1 \text{ (by inequality (47) and our assumption)} \\ &\implies R_{S_2}^{T_1} \geq 1 + 2\epsilon_1 - \epsilon \geq 1 \text{ (since } R_{S_3}^{T_1} \leq \epsilon). \end{aligned} \quad (51)$$

By the definition of separability there needs to exist a subset  $E_{S_2}^{T_1}$  of the edges connecting  $S_2'$  and  $S_2$  such that  $|E_{S_2}^{T_1}| \geq nR_{S_2}^{T_1} \geq n$  and such that the received bits at  $T_1$  uniquely determine the bits in  $E_{S_2}^{T_1}$ . This means that

$$\text{rank}(B_2) \geq n. \quad (52)$$

The matrix  $Z_{T_2}$  can be written as

$$Z_{T_2} = \begin{bmatrix} B_1 & B_2 & 0 \\ 0 & 0 & D \end{bmatrix}. \quad (53)$$

By the constraints of  $\mathcal{R}_{SV}$  we also require that  $R_{S_1}^{T_2} \geq 1$ . For separability to hold, we need the existence of a subset  $E_{S_1}^{T_2}$  of the bits from  $S_1'$  to  $S_1$  so that  $|E_{S_1}^{T_2}| \geq nR_{S_1}^{T_2} \geq n$  and such that the received bits at  $T_2$  uniquely determine the bits in  $E_{S_1}^{T_2}$ . This means that  $\text{rank}(B_1) = n$  and that  $T_2$  needs to receive all of the bits corresponding to the columns of  $B_1$ . By Lemma 3 in the Appendix, we know that if

$$\text{rank} \begin{bmatrix} B_1 \\ 0 \end{bmatrix} + \text{rank} \begin{bmatrix} B_2 & 0 \\ 0 & D \end{bmatrix} > \text{rank} \begin{bmatrix} B_1 & B_2 & 0 \\ 0 & 0 & D \end{bmatrix} \quad (54)$$

then there exist input vectors  $u_x = [u_{1x}^T \ u_{2x}^T \ u_{3x}^T]^T$  and  $u_y = [u_{1y}^T \ u_{2y}^T \ u_{3y}^T]^T$  such that  $u_{1x} \neq u_{1y}$  but  $Z_{T_2}u_x = Z_{T_2}u_y$ . Since the solution is assumed to be separable, this cannot be the case. It is easy to see that we need

$$\begin{aligned} \text{rank} \begin{bmatrix} B_1 & B_2 & 0 \end{bmatrix} &= \text{rank}(B_1) + \text{rank} \begin{bmatrix} B_2 & 0 \end{bmatrix} \\ &= \text{rank}(B_1) + \text{rank}(B_2) \\ &\geq 2n. \end{aligned} \quad (55)$$

However, we also know that  $\text{rank} \begin{bmatrix} B_1 & B_2 & 0 \end{bmatrix} \leq n(1+\epsilon)$  (the rank of a matrix cannot be more than the number of rows). Thus, if  $\epsilon < 1$ , then separability cannot hold at terminal  $T_2$ .  $\square$

It is important to note that counterexamples for higher numbers of sources and receivers can be constructed by simply choosing the counterexamples demonstrated as appropriate subgraphs in the network.

The results mean that one can always find nonseparable distributed source coding problems when either the number of sources or the number of terminals (or both) is larger than two.

While not presented here, similar conclusions (with similar proofs) hold in the cost version of the problem where the edges have cost associated with usage per bit rather than capacity constraints [22].

#### IV. CONCLUSION

The problem of distributed source coding of multiple sources over a network with multiple receivers was considered. In particular, we focused on investigating whether source coding can be separated from the problem of transmitting an appropriate number of coded bits to each receiver. Networks with capacities on edges were considered. While in general the answer is negative, we showed the surprising result that in the specific case of two sources and two receivers, a separable solution always exists.

#### APPENDIX

*Lemma 2:* Consider two matrices  $A$  and  $B$  over  $GF(2)$  of same dimensions. Then

$$\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B) \quad (56)$$

$$\text{rank}(A+B) \geq \text{rank}(A) - \text{rank}(B). \quad (57)$$

*Proof:* The rank of  $(A+B)$  is the number of linearly independent columns in  $(A+B)$ . Therefore  $\text{rank}(A+B) \leq \text{rank}[(A+B)|B]$ . Now

$$\begin{aligned} \text{rank}[(A+B)|B] &= \text{rank} \left( [(A+B)|B] \times \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \right) \\ &\quad \text{(multiplication by an invertible matrix)} \\ &= \text{rank}[A|B]. \end{aligned} \quad (58)$$

Since  $\text{rank}[A|B] \leq \text{rank}(A) + \text{rank}(B)$  we have the proof. To obtain the second part of the lemma we simply apply the first part to matrices  $B$  and  $(A+B)$ .  $\square$

*Lemma 3:* Consider a matrix  $A = [A_1|A_2]$  such that  $\text{dim}(A_1) = m \times n_1$ ,  $\text{dim}(A_2) = m \times n_2$ , and  $\text{rank}(A_1) = n_1$ . If

$$\text{rank}(A_1) + \text{rank}(A_2) > \text{rank}(A)$$

then there exist two vectors  $u_x = \begin{bmatrix} u_{1x} \\ u_{2x} \end{bmatrix}$  and  $u_y = \begin{bmatrix} u_{1y} \\ u_{2y} \end{bmatrix}$  with  $\text{dim}(u_{1x}) = \text{dim}(u_{1y}) = n_1 \times 1$  and  $\text{dim}(u_{2x}) = \text{dim}(u_{2y}) = n_2 \times 1$  such that  $Au_x = Au_y$  and  $u_{1x} \neq u_{1y}$ . In other words, if  $Au_x = Au_y$  implies  $u_x = u_y$ , then  $\text{rank}(A_1) + \text{rank}(A_2) = \text{rank}(A)$ .

*Proof:* We denote the  $\text{rank}(A_2)$  by  $r_2$ . Then we can find a set of  $r_2$  linearly independent columns of  $A_2$  that we denote by  $[c_1|c_2|\dots|c_{r_2}]$ . Consider a matrix  $B$

$$B = [A_1 \ c_1 \ c_2 \ \dots \ c_{r_2}] \quad (59)$$

such that  $\text{rank}(B) \leq \text{rank}(A)$  and the number of columns in  $B = n_1 + r_2$ . Since  $n_1 + r_2 > \text{rank}(B)$  some column in  $B$  is linearly dependent on the others. This means that there exists a nonzero vector  $u^* = \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix}$  such that

$$B \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix} = 0. \quad (60)$$

Since the columns  $[c_1|c_2|\dots|c_{r_2}]$  are linearly independent we know that  $u_1^* \neq 0$ , and since the columns of  $A_1$  are linearly independent we know that  $u_2^* \neq 0$ . Using  $u^*$  we can construct a new nonzero vector  $v = \begin{bmatrix} v_{1x} \\ v_{1y} \end{bmatrix}$  where  $\text{dim}(v_{1x}) = n_1 \times 1$  and  $\text{dim}(v_{2x}) = n_2 \times 1$  such that  $Av = 0$ . To see this note that we can set  $v_{1x} = u_1^*$  and the values of elements in  $v_{1y}$  corresponding to the columns of  $[c_1|c_2|\dots|c_{r_2}]$  equal to the corresponding values of the elements of  $u_2^*$  and zero elsewhere.

Now any two vectors  $u_x$  and  $u_y$  such that  $u_x - u_y = v$  will satisfy  $Au_x = Au_y$  with  $u_{1x} \neq u_{1y}$ . We know that  $\text{rank}(A_1) + \text{rank}(A_2)$  is always greater than or equal to  $\text{rank}(A)$ . Therefore

$$\text{rank}(A_1) + \text{rank}(A_2) = \text{rank}(A) \quad (61)$$

is a necessary condition for the nonexistence of vectors  $u_x$  and  $u_y$  such that  $u_x \neq u_y$  and  $Au_x = Au_y$ .  $\square$

#### ACKNOWLEDGMENT

The authors would like to acknowledge the anonymous reviewers whose comments greatly improved the quality of the correspondence. The first author would like to acknowledge interesting comments from Dr. N. Vaswani and Dr. P. Chaichanavong on this correspondence.

#### REFERENCES

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

- [2] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 4, pp. 585–592, Jul. 1982.
- [3] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense sensor network," *IEEE Signal Process. Mag.*, Mar. 2003.
- [4] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Commun. Lett.*, vol. 5, pp. 417–419, Oct. 2001.
- [5] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, Sep. 2004.
- [6] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [7] S.-Y. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Info. Theory*, vol. 49, no. 2, pp. 371–381, 2003.
- [8] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, 2003.
- [9] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in *Proc. 41st Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2003.
- [10] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. 41st Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2003.
- [11] A. Ramamoorthy, J. Shi, and R. D. Wesel, "On the capacity of network coding for random networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2878–2885, Aug. 2005.
- [12] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked Slepian-Wolf: Theory, algorithms and scaling laws," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4057–4073, Dec. 2005.
- [13] T. Ho, M. Médard, M. Effros, and R. Koetter, "Network coding for correlated sources," in *Proc. Conf. Information Science and Systems*, 2004.
- [14] 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 *Proc. DIMACS Workshop Networking Information Theory*, Piscataway, NJ, 2003.
- [15] A. Wyner, "Recent results in Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 2–10, Jan. 1974.
- [16] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.
- [17] 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.
- [18] T. S. Han, "Slepian-Wolf-Cover theorem for a network of channels," *Inf. Contr.*, vol. 47, no. 1, pp. 67–83, 1980.
- [19] Y. Wu, V. Stanković, Z. Xiong, and S. Y. Kung, "On practical design for joint distributed source coding and network coding," in *Proc. 1st Workshop Netw. Coding, Theory Applicat.*, Riva del Garda, Italy, 2005.
- [20] K. Jain, M. Mahdian, and M. R. Salavatipour, "Packing Steiner trees," in *SODA*, 2003.
- [21] C. Fragouli and E. Soljanin, "Subtree decomposition for network coding," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 145.
- [22] A. Ramamoorthy, K. Jain, P. A. Chou, and M. Effros, "Separating distributed source coding from network coding," Tech. Rep., MSR, to be published.

## Cycle-Logical Treatment for "Cyclopathic" Networks

Ángela I. Barbero and Øyvind Ytrehus, *Senior Member, IEEE*

**Abstract**—This correspondence addresses the problem of finding the network encoding equations for error-free networks with multiple sources and sinks. Previous algorithms could not cope with cyclic networks. Networks that are cyclic in three different senses are considered in this correspondence, and two extensions of the polynomial time Linear Information Flow (LIF) algorithm are presented. The first algorithm will produce the network encoding equations for a network which can be cyclic, unless the actual flow paths form cycles. The second algorithm will work also when the flow paths form simple cycles. Finally an example of a third kind of cyclic network, where the previous algorithms will fail, is given. However, a binary encoding is provided also in this case.

**Index Terms**—Cyclic networks, multicasting, network codes, pipelining.

### I. INTRODUCTION

Consider an error-free multicasting network for transmission of information that is generated and sent by multiple *sources*, forwarded by *routers*, and received by multiple information *sinks*. Ahlswede *et al.* [1] demonstrated that a multiple-sink network where routers can perform network coding can offer a higher overall throughput than one where the routers simply forward the information messages.

Several approaches have been followed in order to design an algorithm that produces the encoding equations used by the routers. For example, random, distributed and time variant encoding with symbols from a large field (see [2], [3]) can be very robust and requires no a priori knowledge of the network structure or the flow paths used, but this induces an overhead since the encoding equations must be transmitted with each individual information message.

We consider the problem of designing a centralized algorithm for calculating the network encoding equations, and in particular in the case of networks that contain cyclic paths (which we call *cyclopathic networks*.)

Sanders *et al.* [4] and Jaggi *et al.* [5] presented the efficient Linear Information Flow (LIF) algorithm for producing the network encoding equations. As shown by Ahlswede *et al.* [1], network coding can be applied also to cyclic networks. The LIF algorithm will not, however, work in the case where the network contains cycles. Fragouli and Soljanin [6] provide an interesting notation for describing the encoding equations, also for networks that contain cycles, but do not relate the code design process for cyclic networks to the efficient approach of the LIF algorithm.

In Section II, we give an overview of the necessary notation and previous results. Further, we will define two different notions of cyclicity in networks: We will consider *link cyclic* networks as well as the more difficult case of *flow cyclic* networks. Next, we restate the LIF algorithm as an edge following algorithm. The new version, the LIFE algorithm, is presented in Section III, and it is shown that the LIFE algorithm works on networks that are link cyclic, but flow acyclic. However,

---

Manuscript received March 13, 2005; revised January 28, 2006. The work of Á. I. Barbero was supported by grant PR2005-0011 of Secretaría de Estado de Universidades e Investigación, Ministerio de Educación y Ciencia (Spanish Government). The work of Ø. Ytrehus was supported by the Norwegian Research Council (NFR).

Á. I. Barbero is with the Department of Applied Mathematics, University of Valladolid, 47011 Valladolid, Spain (e-mail: angbar@wmatem.eis.uva.es).

Ø. Ytrehus is with the Department of Informatics, University of Bergen, N-5020 Bergen, Norway (e-mail: oyvind@ii.uib.no).

Communicated by N. Cai, Guest Editor.

Digital Object Identifier 10.1109/TIT.2006.874542