On Networks with Side Information

Asaf Cohen, Salman Avestimehr and Michelle Effros

Department of Electrical Engineering

California Institute of Technology

Pasadena, CA 91125, USA

{soofsoof, avestime, effros}@caltech.edu

Abstract—In this paper, we generalize the lossless coded side information problem from the three-node network of Ahlswede and Körner to more general network scenarios. We derive inner and outer bounds on the achievable rate region in the general network scenario and show that they are tight for some families of networks. Our approach demonstrates how solutions to canonical source coding problems can be used to derive bounds for more complex networks and reveals an interesting connection between networks with side information, successive refinement, and network coding.

I. INTRODUCTION

We consider a lossless source coding problem over noiseless links. Sources X and Y are available at distinct nodes in the network. Only source X is required at one or more sink nodes. The resulting family of coded side information problems captures scenarios where the network capacity is insufficient to describe a source X to its intended destinations; in such scenarios, X can sometimes still be delivered without loss provided sufficient capacity from a helper random variable Y. We wish to characterize this family of scenarios. When Y is independent of X, our coded side information problem reduces to the multicast problem [1]. When X and Y are dependent and Y is available only at a sink node from which there are no outgoing edges, the problem reduces to a multicast problem with side information [2]. When X and Y are dependent and the network is the three-node network with independent transmitters observing X and Y and a single receiver reconstructing X, this is the original coded side information problem [3]. While the rate region is solved for each of the special cases above, the general problem remains unsolved. Network models with more than two sources were considered in [4] and [5], however, in these models each node in the network is either a source or a sink, hence they do not include the scenarios we consider here.

At the basis of key results in source coding for networks, e.g. [1], [6] and [2], stands a cut-set argument, namely, a partition of the network into two disjoint sets and the usage of known results from source coding for a point-to-point link. Yet, it is clear that even for small networks such an argument can be loose, and the question that arises is how can we harness solutions to canonical problems more complex than point-to-point in order to bound the rate region for large networks.

In this paper, we derive inner and outer bounds on the rate region for networks with side information when the side information is available at a non-sink node. In particular, we show how the solutions to source coding problems like the coded side information problem of [3] or the partially separated encoders problem of [7] can be used to bound the rate region of large networks with dependent sources. We describe sufficient conditions for the tightness of our bounds and show their relation to the conditions for successive refinability in lossy source coding [8].

The rest of the paper is organized as follows. Section II includes a formal statement of the problem. In Section III, we derive inner and outer bounds on the set of achievable rates. In Section IV, we give sufficient conditions for the tightness of our bounds and demonstrate a few network scenarios where those conditions apply. Section V includes a few simple examples.

II. PROBLEM STATEMENT

The following definitions are used throughout the paper. A network is defined as a directed acyclic graph $(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices (nodes) and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges (links). For each edge $e = (a, b) \in \mathcal{E}$, we use $o(e) = a$ and $d(e) = b$ to denote the origin and destination vertices, respectively, of edge e. Associated with each edge $e \in \mathcal{E}$ is a capacity $c(e) \geq 0$. We assume noise-free links.

We specify a *side information network* $(\mathcal{V}, \mathcal{E}, s, z, T)$ as follows. Let $\{(X_i, Y_i)\}_{i=1}^{\infty}$ be a sequence of independent and identically distributed pairs of discrete random variables with alphabet $\mathcal{X} \times \mathcal{Y}$. Source $\{X_i\}$ is available at node $s \in \mathcal{V}$, side information $\{Y_i\}$ is available at node $z \in \mathcal{V} \setminus \{s\}$, and source $\{X_i\}$ is demanded losslessly at the sink nodes $T \subseteq \mathcal{V} \setminus \{s, z\}$. We assume that nodes s and z have no incoming edges and that each sink node in T has no outgoing edges.

Let $\mathcal{E}_{XY} \subseteq \mathcal{E}$ denote the set of edges for which there is a directed path from s to $o(e)$ with a strictly positive capacity, and denote by $\mathcal{E}_Y$ the set $\mathcal{E} \setminus \mathcal{E}_{XY}$. Given any non-intersecting sets $A, B \subset \mathcal{V}$, a cut $\mathcal{V}_{A;B}$ is a subset of vertices which includes A but is disjoint from B, that is, $A \subseteq \mathcal{V}_{A;B}$ and $\mathcal{V}_{A;B} \cap B = \emptyset$. Let $\mathcal{C}(\mathcal{V}_{A;B})$ be the set of edges $e \in \mathcal{E}$ for which $o(e) \in \mathcal{V}_{A;B}$ and $d(e) \notin \mathcal{V}_{A;B}$ and $\mathcal{V}_{A;B}^*$ describe the cut for which $\sum_{e \in \mathcal{C}(\mathcal{V}_{A;B})} c(e)$ is minimal among all cuts $\mathcal{V}_{A;B}$. For example, $\mathcal{V}_{s,z;t}$ describes a cut that separates source nodes

$\{s, z\}$ from receiver $t \in T$, $\mathcal{C}(\mathcal{V}_{s,z;t})$ is the set of edges that traverse that cut, and $\mathcal{V}_{s,z;t}^*$ is the min-cut between $\{s, z\}$ and $\{t\}$. (We omit the brackets from $\{s, z\}$ and $\{t\}$ for notational concision.)

For any vector of rates $(R_e)_{e \in \mathcal{E}}$, a $((2^{nR_e})_{e \in \mathcal{E}}, n)$ source code comprises the following mappings

$$
\begin{aligned}
g_e^n &: \mathcal{X}^n \mapsto \{1, \ldots, 2^{nR_e}\} & e \in \mathcal{E}, o(e) = s \\
g_e^n &: \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_e}\} & e \in \mathcal{E}, o(e) = z \\
g_e^n &: \Pi_{e':d(e')=o(e)} \{1, \ldots, 2^{nR_{e'}}\} \\
&\mapsto \{1, \ldots, 2^{nR_e}\} & e \in \mathcal{E}, o(e) \notin \{s, z\} \\
h_t^n &: \Pi_{e:d(e)=t} \{1, \ldots, 2^{nR_e}\} \mapsto \mathcal{X}^n & t \in T.
\end{aligned}
$$

For each $t \in T$, we use $\hat{X}_t^n$ to denote the reproduction of X^n found by decoder h_t^n. We are interested in the set of possible values $(c(e))_{e \in \mathcal{E}}$ for which for any $\epsilon > 0$ there exists a sufficiently large n and a $((2^{nR_e})_{e \in \mathcal{E}}, n)$ code with $R_e \leq c(e)$ for all $e \in \mathcal{E}$, such that $\Pr(\hat{X}_t^n = X^n) \geq 1 - \epsilon$ for all sinks $t \in T$. That is, X can be reproduced at the terminals $t \in T$ with a negligible error probability. We call the closure of this set of rate vectors the *set of achievable rates*, which we denote by $\mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, T)$. In the rest of the paper, we derive general inner and outer bounds on this set and investigate the scenarios in which these bounds are tight.

III. OUTER AND INNER BOUNDS ON THE SET OF ACHIEVABLE RATES

In this section, we derive inner and outer bounds on the set of achievable rates for side information networks. The following outer bound on the set of achievable rates generalizes the cut-set bound [9, Theorem 14.10.1] to give a necessary condition for lossless reconstruction of X at the terminal nodes given side information Y at any other node in the network.

Theorem 1. *Given a side information network $(\mathcal{V}, \mathcal{E}, s, z, T)$, if $(c(e) : e \in \mathcal{E}) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, T)$, then for each $t \in T$ and each cut $\mathcal{V}_{s,z;t}$ there exists a random variable $U \in \mathcal{U}$ such that $U \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}| \leq |\mathcal{Y}|$, and*

$$
\begin{aligned}
\sum_{e \in \mathcal{E}_{XY} \cap \mathcal{C}(\mathcal{V}_{s,z;t})} c(e) &\geq H(X|U) \\
\sum_{e \in \mathcal{E}_Y \cap \mathcal{C}(\mathcal{V}_{s,z;t})} c(e) &\geq I(Y; U).
\end{aligned}
$$

In Section IV we give sufficient conditions on the network for which Theorem 1 is tight. While the region is not tight in general, it is always at least as strong as the cut-set bound. To see this, note that for any cut $\mathcal{V}_{s,z;t}$, the cut-set bound is $\sum_{e \in \mathcal{C}(\mathcal{V}_{s,z;t})} c(e) \geq H(X)$ while Theorem 1 gives

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{s,z;t})} c(e) \geq H(X|U) + I(Y; U).
$$

Since $H(X|U) + I(Y; U) \geq H(X)$ for all $U \leftrightarrow Y \leftrightarrow X$, the second bound is at least as good as the first. On the other hand, for any cut $\mathcal{V}_{s;z,t}$, the cut-set bound is $\sum_{e \in \mathcal{C}(\mathcal{V}_{s;z,t})} c(e) \geq H(X|Y)$ while taking $\tilde{\mathcal{V}}_{s,z;t} = \mathcal{V}_{s;z,t} \cup \{z\}$ and using Theorem 1 and the observation that all edges leaving node z lie in
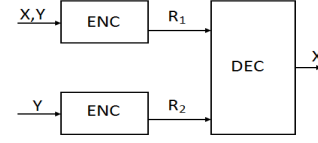

Fig. 1. Coded side information with Y known at the X encoder.

$\mathcal{E}_Y$ rather than $\mathcal{E}_{XY}$ gives

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{s;z,t})} c(e) \geq \sum_{e \in \mathcal{E}_{XY} \cap \mathcal{C}(\tilde{\mathcal{V}}_{s,z;t})} c(e) \geq H(X|U),
$$

which is again at least as tight as the cut-set bound.

Before proving Theorem 1, we describe the rate region for the network shown in Figure 1. While the network differs from the coded side information problem in that side information Y is available to the encoder of X, the rate region is the same as the rate region for the coded side information problem by a converse similar to the one in [3], which we omit for brevity. The proof that $|\mathcal{U}| \leq |\mathcal{Y}|$ follows from [10, Theorem 1].

Proposition 2. *Consider the source coding problem given in Figure 1. Rate vector (R_1, R_2) is achievable if and only if there exists a random variable $U \in \mathcal{U}$ such that $U \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}| \leq |\mathcal{Y}|$, and*

$$
R_1 \geq H(X|U); \quad R_2 \geq I(Y; U).
$$

Theorem 1 Proof: Let $\{((g_e^n)_{e \in \mathcal{E}}, (h_t^n)_{t \in T})\}_{n=1}^{\infty}$ be a sequence of $((2^{nc(e)})_{e \in \mathcal{E}}, n)$ codes for which $\Pr(\hat{X}_t^n = X^n) \to 1$ as $n \to \infty$ for all $t \in T$. Given any receiver $t \in T$ and any cut $\mathcal{V}_{s,z;t}$, let $f_e^n(X^n, Y^n)$ denote the message carried along edge e for each $e \in \mathcal{E}_{XY} \cap \mathcal{C}(\mathcal{V}_{s,z;t})$ and $g_e^n(Y^n)$ denote the message carried along edge e for each $e \in \mathcal{E}_Y \cap \mathcal{C}(\mathcal{V}_{s,z;t})$. Further, define the vectors

$$
\begin{aligned}
f^n(X^n, Y^n) &= (f_e^n(X^n, Y^n))_{e \in \mathcal{E}_{XY} \cap \mathcal{C}(\mathcal{V}_{s,z;t})} \\
g^n(Y^n) &= (g_e^n(Y^n))_{e \in \mathcal{E}_Y \cap \mathcal{C}(\mathcal{V}_{s,z;t})}
\end{aligned}
$$

Then (f^n, g^n, h_t^n) can be viewed as a code of rate $(\sum_{e \in \mathcal{E}_{XY} \cap \mathcal{C}(\mathcal{V}_{s,z;t})} c(e), \sum_{e \in \mathcal{E}_Y \cap \mathcal{C}(\mathcal{V}_{s,z;t})} c(e))$ for the network in Figure 1, giving the desired result by Proposition 2. ∎

We now derive an inner bound on the set of achievable rates for networks with one source, one destination, and one side information node.

Lemma 3. *Let $(\mathcal{V}, \mathcal{E}, s, z, \{t\})$ be a side information network. If there exists a vertex $v \in \mathcal{V}$ and a random variable $U \in \mathcal{U}$ such that $U \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}| \leq |\mathcal{Y}|$, and*

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{s;v}^*)} c(e) \geq H(X|U) \tag{1}
$$

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{z;v}^*)} c(e) \geq I(Y; U) \tag{2}
$$

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{s,z;v}^*)} c(e) \geq H(X|U) + I(Y; U) \tag{3}
$$

$$
\sum_{e \in \mathcal{C}(\mathcal{V}_{v;t}^*)} c(e) \geq H(X), \tag{4}
$$
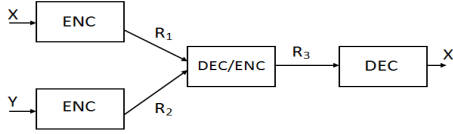
Fig. 2. A simple 4-node network.

Fig. 3. A network with k demands, direct links from Y.

Fig. 4. A network with two demands, arbitrary Y network.

then $(c(e))_{e\in\mathcal{E}} \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t\})$.

Lemma 3 is proved in [11]. The essence of the proof is the fact that at the rate given in Lemma 3 a simple multiple-unicast code can send the required descriptions of X to v, reconstruct it and forward the reconstruction to t. Figure 2 includes a simple example for which the bound in Lemma 3 improves on the achievability result from [2, Theorem 3].

Both Theorem 1 and Lemma 3 demonstrate the power of small, canonical problems, in the solution of large networks. Any outer bound for a canonical problem yields a corresponding outer bound for general networks of the corresponding structure. Any known solution to such a problem yields an achievable region. Moreover, additional properties of the codes, such as error exponents, are likewise inherited. For example, it is clear from the proof of Lemma 3 that linear encoding and decoding can be used at all nodes but z and v. Finally, the result in Lemma 3 can be generalized to the cases where the intermediate node v, destination t, or both have additional side information available, using the generalization of [3] given in [12].

## IV. SUFFICIENT CONDITIONS FOR THE TIGHTNESS OF THEOREM 1

In this section, we derive sufficient conditions under which the inner and outer bounds of Section III are tight. Corollary 4 treats networks where the only node with directed paths from both the source and side information nodes is the sink.

Corollary 4. *Let* $(\mathcal{V}, \mathcal{E}, s, z, \{t\})$ *be a side information network. If* t *is the only vertex in* $\mathcal{V}$ *for which there are directed paths from both* s *and* z, *then Theorem 1 is tight and we can rewrite the achievable rate region as* $(c(e) : e \in \mathcal{E}) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t\})$ *if and only if there exists a random variable* $U \in \mathcal{U}$ *such that* $U \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}| \leq |\mathcal{Y}|$, *and*

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{s;t}^*)} c(e) \geq H(X|U)$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{z;t}^*)} c(e) \geq I(Y;U).$$

Proof: We prove the achievability of Theorem 1 using Lemma 3. Note first that any cut $\mathcal{V}_{s,z;t}$ can be written as a union of cuts $\mathcal{V}_{s;t}$ and $\mathcal{V}_{z;t}$ for which $\mathcal{C}(\mathcal{V}_{s,z;t} \cap \mathcal{E}_{XY}) = \mathcal{C}(\mathcal{V}_{s;t})$ and $\mathcal{C}(\mathcal{V}_{s,z;t} \cap \mathcal{E}_Y) = \mathcal{C}(\mathcal{V}_{z;t})$, which gives the rate region characterization above. This gives us (1) and (2) immediately. Condition (4) is also trivially satisfied. It remains
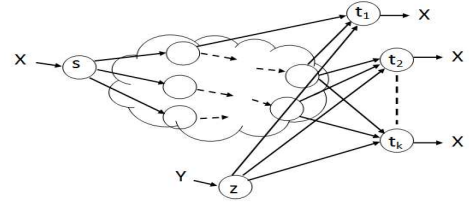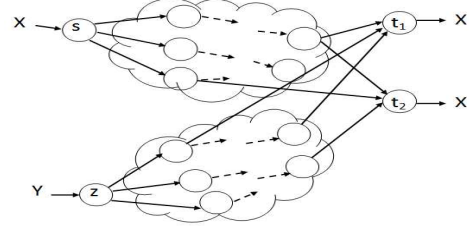
to check condition (3). For any cut $\mathcal{V}_{s,z;t}$,

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{s,z;t})} c(e) \geq \sum_{e \in \mathcal{C}(\mathcal{V}_{s;t}^*)} c(e) + \sum_{e \in \mathcal{C}(\mathcal{V}_{z;t}^*)} c(e) \quad (5)$$

$$\geq H(X|U) + I(Y;U)$$

where (5) follows since $\mathcal{C}(\mathcal{V}_{s;t}^*) \cap \mathcal{C}(\mathcal{V}_{z;t}^*) = \emptyset$. ∎

We next consider networks with k sinks ($T = \{t_1, \ldots, t_k\}$). Theorem 5 treats the case where $\mathcal{E}_Y = \{(z, t) : t \in T\}$; that is, there is a direct link from z to t_i and that is the only path from z to t_i for each $t_i \in T$, as shown in Figure 3. Clearly, Theorem 5 will continue to hold if each of the links in $\mathcal{E}_Y$ is replaced with a separate network of equal capacity. The details are omitted.

Theorem 5. *Let* $(\mathcal{V}, \mathcal{E}, s, z, \{t_1, \ldots, t_k\})$ *be a side information network with* $\mathcal{E}_Y = \{(z, t_i)\}_{i=1}^k$. *Then Theorem 1 is tight, and we can rewrite the resulting rate region as* $(c(e) : e \in \mathcal{E}) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t_1, \ldots, t_k\})$ *if and only if there exist random variables* $U_i \in \mathcal{U}_i$, $i = 1, \ldots, k$ *such that for each* i, $U_i \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}_i| \leq |\mathcal{Y}|$, *and*

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{s;t_i}^*)} c(e) \geq H(X|U_i) \quad (6)$$

$$c((z, t_i)) \geq I(Y;U_i) \quad (7)$$

Proof sketch: By [3] and [10, Theorem 1], the existence of U_i, $i = 1, \ldots, k$, satisfying the given conditions implies that for any $i = 1, \ldots, k$, $\epsilon, \delta > 0$ and n sufficiently large there exist encoding functions $f_i^n : \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_Y^i}\}$ and $g_i^n : \mathcal{X}^n \mapsto \{1, \ldots, 2^{nR_X^i}\}$ and decoding function $h_i^n : \{1, \ldots, 2^{nR_Y^i}\} \times \{1, \ldots, 2^{nR_X^i}\} \mapsto \mathcal{X}^n$ satisfying

$$\Pr(\hat{X}_{t_i}^n = X^n) \geq 1 - \epsilon$$
$$R_X^i \leq H(X|U_i) + \delta$$
$$R_Y^i \leq I(Y;U_i) + \delta.$$

1345

Hence, by (7), the rates on links (z, t_i), $i = 1, \ldots, k$, suffice for the transmission of the $f_i^n(Y^n)$ values.

As for the $g_i^n(X^n)$ values, note that for general functions $g_i^n(\cdot)$ of cardinality $2^{nH(X|U_i)}$, satisfying (6) alone does not suffice to communicate these values to their corresponding destinations. However, since the g_i^n functions can be implemented using random binning, the following scheme can be used. Randomly bin the x^n sequences to $2^{nH(X|U_{i^*})}$ bins, where i^* maximizes $H(X|U_i)$. Use a *linear broadcast code*, in the sense of [13, Definition 2.9], to send the $H(X|U_{i^*})$ bits that represent the bin index of the received x^n sequence across the X network. That is, each terminal t_i receives m_{t_i} independent equations on the $H(X|U_{i^*})$ bits, where $m_{t_i} = \min\{\sum_{e \in \mathcal{C}(\mathcal{V}^*_{s;t_i})} c(e), H(X|U_{i^*})\}$ ([13, Corollary 2.23]). The null-space dimension of a terminal t_i is thus $H(X|U_{i^*}) - m_{t_i}$, resulting in an effective random bin of size $2^{nH(X|U_i)}$ which includes the true x^n. This, together with $f_i^n(y^n)$, suffices for the reconstruction of x^n. ∎

Theorem 6 treats networks with an arbitrary set of connections between the side information node and the sinks $T = \{t_1, t_2\}$ as shown in Figure 4 when X and Y are uniformly distributed binary random variables related through a binary symmetric channel (BSC).

Theorem 6. *Consider a side information network $(\mathcal{V}, \mathcal{E}, s, z, \{t_1, t_2\})$ where both the source and the side information are binary symmetric random variables. If the only vertices $v \in \mathcal{V}$ for which there is a directed path from both s and z are t_1 and t_2 (see Figure 4), then Theorem 1 is tight, and we can rewrite the resulting rate region as $(c(e) : e \in \mathcal{E}) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t_1, t_2\})$ if and only if there exist random variables $U_1 \in \mathcal{U}_1$ and $U_2 \in \mathcal{U}_2$ such that $U_1 \leftrightarrow Y \leftrightarrow X$, $U_2 \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}_1| \leq |\mathcal{Y}|$, $|\mathcal{U}_2| \leq |\mathcal{Y}|$, and*

$$\sum_{e \in \mathcal{C}(\mathcal{V}^*_{s;t_1})} c(e) \geq H(X|U_1) \tag{8}$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}^*_{z;t_1})} c(e) \geq I(Y;U_1) \tag{9}$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}^*_{s;t_2})} c(e) \geq H(X|U_2) \tag{10}$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}^*_{z;t_2})} c(e) \geq I(Y;U_2) \tag{11}$$

Proof sketch: The proof of Theorem 5 shows that conditions (8) and (10) suffice in order to multicast the X descriptions to the sinks. We next show that conditions (9) and (11) suffice in order to multicast the Y descriptions to the sinks when X and Y are binary symmetric random variables related through a BSC.

In [10], the authors have analyzed the problem of source coding with coded side information for binary sources, and have shown that the optimal auxiliary random variable U is also binary, and related to Y through a BSC. Since any two memoryless symmetric binary sources can be described as related through a BSC, for any two auxiliary binary random

variables U_1 and U_2 we can construct $\tilde{U}_1$ and $\tilde{U}_2$ that have the same conditional distributions as U_1 and U_2 yet also satisfy $\tilde{U}_2 \leftrightarrow \tilde{U}_1 \leftrightarrow Y \leftrightarrow X$ or $\tilde{U}_1 \leftrightarrow \tilde{U}_2 \leftrightarrow Y \leftrightarrow X$ (the details are in [11]). Without loss of generality, we assume the former, and supress the tilde notation in the reminder of the proof.

To complete the proof we need to show a construction of the Y descriptions at rates $I(Y;U_2)$ and $I(Y;U_1) = I(Y;U_2) + I(Y;U_1|U_2)$, such that the latter is a refined description of the former. If such a nested structure of descriptions is achieved, by [14, Corollary 2] the two descriptions can be communicated to the sinks if the associated min-cut bounds are satisfied, which in turn holds by conditions (9) and (11).

The description at rate $I(Y;U_2)$ is similar to the original coded side information problem. Simply generate $2^{nI(Y;U_2)}$ sequences, $\{U_2^n(s)\}$, at random according to $\Pi_{i=1}^n p_{U_2}(u_2)$. The coarse description is the index s of a $U_2^n(s)$ such that $(Y^n, U_2^n(s)) \in A_\epsilon^{(n)}(Y, U_2)$, where $A_\epsilon^{(n)}(Y, U_2)$ is the set of (y^n, u_2^n) sequences that are jointly typical. As for the refined description, for each $u_2^n(s)$, generate $2^{nI(Y;U_1|U_2)}$ sequences, $\{U_1^n(s,k)\}$, at random according to $\Pi_{i=1}^n p_{U_1|U_2}(u_1|u_{2,i}(s))$. The refinement description is the index k of a $U_1^n(s,k)$ such that $(Y^n, U_1^n(s,k)) \in A_\epsilon^{(n)}(Y, U_1)$. Due to space limitations, the detailed error analysis is omitted.

∎

Remark 1. *Note that Theorem 6 holds for any X and Y for which the desired auxiliary random variables U_1 and U_2 satisfy $U_2 \leftrightarrow U_1 \leftrightarrow Y \leftrightarrow X$ (or can be constructed this way). The key here is that the description of X is always successively refinable, but the successive refinement property need not hold for all descriptions of Y since these descriptions are lossy.*

Remark 2. *A key finding in [1], [6] and [2] is the ability to analyze the rate region of multicast networks (with independent and dependent sources) only through the values of the min-cuts between sources and destinations. The results given in Corollary 4, Theorem 5, and Theorem 6 extend the range of network scenarios where this min-cut analysis holds, yet using a generalization of the cut-set bound [9, Theorem 14.10.1].*

While Theorem 1 gives an outer bound for any network with side information, considering any node with a path from a source as having complete information about that source leads to a loose bound for some networks. In [11] we harness the results of [7] on partially separated encoders to give a result tighter than Theorem 1 for some networks. For example, consider the network in Figure 5. Due to the links from the X network to the Y network, Theorem 6 does not apply. Yet, the following Lemma gives a tighter outer bound than Theorem 1.

Lemma 7. *Let $(\mathcal{V}, \mathcal{E}, s, z, T)$ be a side information network. If $(c(e) : e \in \mathcal{E}) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, T)$, then for any partition of $\mathcal{V}$ into three disjoint sets S, Z and D satisfying the conditions*

1) *S includes the source node s and has no incoming links;*
2) *Z includes the side information node z and has incoming links only from S;*
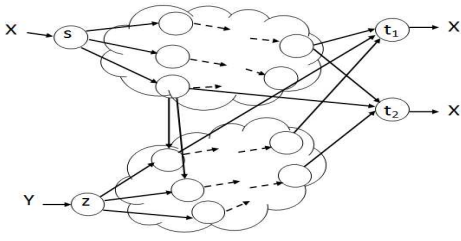3) *$T \subseteq D$;*

1346

Fig. 5. A network with two demands, arbitrary Y network, and links from the X network to the Y network.

the following holds: There exist random variables $W \in \mathcal{W}$, $V \in \mathcal{V}$, $U \in \mathcal{U}$ *of bounded alphabet size and a function* f *such that* $W \leftrightarrow X \leftrightarrow Y$, $U \leftrightarrow (Y, W) \leftrightarrow (X, W) \leftrightarrow V$, $H(X|f(U, V, W)) \leq \epsilon$ *for any* $\epsilon > 0$ *and*

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{S, Z \cup D}):d(e) \in Z} c(e) \geq I(X; W|Y)$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{S, Z \cup D}):d(e) \in D} c(e) \geq I(X; V|U, W)$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{Z, D \cup S})} c(e) \geq I(Y; U|V, W)$$

$$\sum_{e \in \mathcal{C}(\mathcal{V}_{S \cup Z, D})} \geq I(X, Y; U, V, W).$$

V. EXAMPLES

In this section, we illustrate our main results through three simple examples.
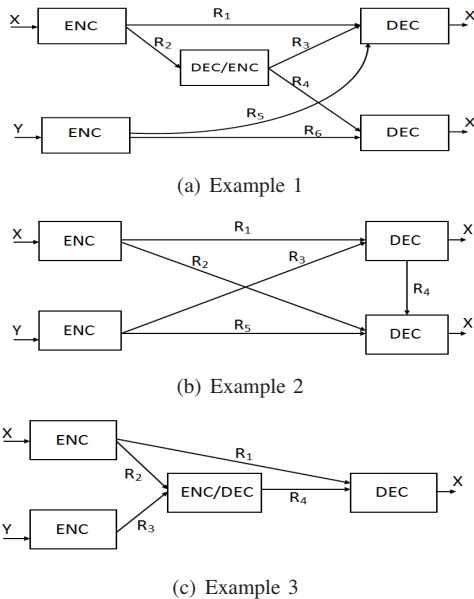

(a) Example 1

(b) Example 2

(c) Example 3

Fig. 6. Network examples.

Example 1. Consider the network in Figure 6(a). By Theorem 5, $(R_1, \ldots, R_6) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t_1, t_2\})$ if and only if there exist random variables $U_1 \in \mathcal{U}_1$ and $U_2 \in \mathcal{U}_2$ such

that for $i \in \{1, 2\}$, $U_i \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}_i| \leq |\mathcal{Y}|$, and

$$R_1 + \min(R_2, R_3) \geq H(X|U_1); \quad R_5 \geq I(Y; U_1)$$
$$\min(R_2, R_4) \geq H(X|U_2); \quad R_6 \geq I(Y; U_2).$$

Example 2. Consider the network in Figure 6(b). By Theorem 1, if $(R_1, \ldots, R_5) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t_1, t_2\})$, then there exist random variables $U_1 \in \mathcal{U}_1$ and $U_2 \in \mathcal{U}_2$ such that $U_1 \leftrightarrow Y \leftrightarrow X$, $U_2 \leftrightarrow Y \leftrightarrow X$, $|\mathcal{U}_1| \leq |\mathcal{Y}|$, $|\mathcal{U}_2| \leq |\mathcal{Y}|$, and

$$R_1 \geq H(X|U_1); \quad R_3 \geq I(Y; U_1)$$
$$R_2 + R_4 \geq H(X|U_2); \quad R_5 \geq I(Y; U_2).$$

While the inner bound in Lemma 3 does not apply here, it is not hard to see that if these conditions are satisfied, $(R_1, \ldots, R_5) \in \mathcal{R}(\mathcal{V}, \mathcal{E}, s, z, \{t_1, t_2\})$.

Example 3. Consider the network in Figure 6(c). While the set of rates $R_1 = H(X|U)$, $R_2 = 0$, $R_3 = I(Y; U)$ and $R_4 = I(X; U)$ with $U \leftrightarrow Y \leftrightarrow X$ satisfies the conditions in Theorem 1, it is clear that these rates are not achievable. This example illustrates why partitioning the network into only two subsets of nodes in the outer bound fails for certain cases, as assuming the ENC/DEC node has X is too restrictive.

ACKNOWLEDGMENT

REFERENCES

[1] R. Ahlswede, N. Cai, S-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.
[2] M. Bakshi and M. Effros, "On achievable rates for multicast in the presence of side information," in *Proc. ISIT*, Toronto, Canada, July 2008, pp. 1661–1665.
[3] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channel," *IEEE Trans. Inform. Theory*, vol. 21, no. 6, pp. 629–637, November 1975.
[4] I. Csiszar and J. Körner, "Towards a general theory of source networks," *IEEE Trans. Inform. Theory*, vol. 26, no. 2, pp. 155–165, March 1980.
[5] T. S. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Trans. Inform. Theory*, vol. 26, no. 3, pp. 277–288, May 1980.
[6] T. Ho, M. Médard, M. Effros, and R. Koetter, "Network coding for correlated sources," in *CISS*, 2004.
[7] A. M. Kaspi and T. Berger, "Rate-distortion for correlated sources with partially separated encoders," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 828–840, Novenber 1982.
[8] B. Rimoldi, "Successive refinement of information: characterization of the achievable rates," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 253–259, January 1994.
[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
[10] W.-H. Gu, R. Koetter, M. Effros, and T. Ho, "On source coding with coded side information for a binary source with binary side information," in *Proc. ISIT*, Nice, France, June 2007, pp. 1456–1460.
[11] A. Cohen, S. Avestimehr, and M. Effros, "On networks with side information," *In preparation*, 2009.
[12] W.-H. Gu and M. Effros, "Simple bounds for lossless source coding in a two-hop network," in *Proc. ITA*, San Diego, California, 2006.
[13] R. Yeung, S.-Y.R. Li, N. Cai, and Z. Zhang, "Network coding theory," *Foundation and Trends in Communications and Information Theory*, vol. 2, no. 4–5, pp. 241–381, 2005.
[14] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. on Networking*, vol. 11, no. 5, pp. 782–794, October 2003.