DIMACS Series in Discrete Mathematics
and Theoretical Computer Science

On Networks of Two-way Channels

Gerhard Kramer and Serap A. Savari

ABSTRACT. A network of two way channels (TWCs) is specified by a graph having an edge between vertex u and vertex v if there is a TWC between these vertices. A bidirected cut-set bound is developed for such networks when network coding is permitted, and some implications of this bound are discussed. For example, the bound generalizes and improves upon a flow cut-set bound that is standard in network optimization theory. It follows that routing is rate-optimal if routing achieves the standard flow cut-set bound. Other consequences are that, for several practical networks of TWCs, linear network coding is optimal for multicasting, and one can separate channel and network coding for multicasting.

Presented at the DIMACS Workshop on Algebraic Coding Theory and Information Theory, DIMACS Center, Rutgers University, Dec. 15-18, 2003.

1. Introduction

Consider a *network* defined by a graph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ with respective vertex and edge sets

$$\mathcal{V} = \{1, 2, \ldots, T\} \tag{1}$$

$$\mathcal{E} = \{(u_1, v_1), (u_2, v_2), \ldots, (u_E, v_E)\} \tag{2}$$

where $u_e, v_e \in \mathcal{V}$ for $e = 1, 2, \ldots, E$. The graph $\mathcal{N}$ might represent a communication, transportation, or distribution network [**AMO93**, § 1.3]. We are interested in *communication* networks where the vertices represent terminals and the edges channels (the word "terminal" here refers generically to any type of vertex, i.e., a terminal need not be the source or sink for a message). We are further interested in channels where *bidirected* transmission is permitted. We propose that the appropriate communication model for this type of scenario is the two-way channel (TWC) introduced in [**S61**]. A TWC representing edge (u, v) is defined by a conditional probability distribution

$$P\left(y_{uv}, y_{vu} | x_{uv}, x_{vu}\right) \tag{3}$$

Key words and phrases. Capacity, coding, information, networks, two-way channel.
The work of G. Kramer was partially supported by the Board of Trustees of the University of Illinois Subaward No. 04-217 under NSF Grant No. CCR-0325673. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University or its Project Director or of the National Science Foundation. The work of S. A. Savari was supported in part by DIMACS.

where x_{uv} and y_{vu} are the respective input to and output from edge (u, v) at vertex u. The terminals can further perform joint channel and *network* coding [**ACLY00, KM03**], i.e., each terminal can transmit into each of its TWCs any function of its messages and past received outputs. We will precisely define this type of operation in the next section, but we begin by giving more details about our network model.

2. Communication Network Model

We consider a model that is a special case of the discrete memoryless network (DMN) in [**K03**] (see also [**vdM68, vdM77, C75**], [**CT91**, Ch. 14]). The DMN has a channel defined by a conditional probability distribution

$$P\left(y_1, y_2, \ldots, y_T | x_1, x_2, \ldots, x_T\right) \tag{4}$$

where x_t and y_t are the respective inputs and outputs of terminal t. The x_t and y_t are therefore realizations of *vector* random variables X_t and Y_t, since each vertex has one channel input and output for every edge incident to it.

The above notation labels the x_t and y_t by their terminals or, for networks of TWCs, by their graph *vertices*. A more convenient approach here will be to label these inputs and outputs by their graph *edges*. We thus define a network of TWCs to be the special case of a DMN where (4) factors as

$$\prod_{(u,v)\in\mathcal{E}} P\left(y_{uv}, y_{vu} | x_{uv}, x_{vu}\right). \tag{5}$$

As in (3), the meaning is that x_{uv} and y_{vu} are the respective input to and output from edge (u, v) at vertex u. The letters x_{uv} and y_{vu} are realizations of the respective random variables X_{uv} and Y_{vu}. The relation of x_{uv} to the x_t of (4) is that the former is an entry of one of the x_t.

Each TWC is characterized by a two-dimensional *capacity region* $\mathcal{C}_{uv}$, i.e., a set of rate pairs (R_{uv}, R_{vu}) that specifies at what rates one can transmit in both directions simultaneously and reliably, where R_{uv} is the rate going from u to v. These rates are characterized in [**S61**, § 15] but they can be difficult to compute.

Consider, for instance, a network with *directed* edges. The TWC capacity regions are then single dimensional, i.e., they consist of rate pairs $(R, 0)$ satisfying $0 \le R \le C_{uv}$ for some non-negative C_{uv} (see Fig. 1). The number C_{uv} is the *capacity* of edge (u, v). This model was considered in [**B02**], and our results extend the findings of [**B02**] to networks of TWCs.

Consider next *undirected* edges. Such edges correspond to half-duplex transmission, and $\mathcal{C}_{uv}$ is the triangle whose vertices are at $(0, 0)$, $(C_{uv}, 0)$, $(0, C_{uv})$. For instance, the *push-to-talk* TWC has such a capacity region [**S61**, Fig. 5]. Moreover, this is an appropriate model for problems where terminals perform time-division multiplexing (one can similarly treat frequency-division duplexing).

Finally, consider channels where $\mathcal{C}_{uv}$ is the square whose extreme point is (C_{uv}, C_{uv}). Such a situation might occur when full-duplex transmission is possible, e.g., a digital subscriber line (DSL) with modems that perform perfect echo cancellation. We can, in fact, replace the edge of such a TWC by a pair of oppositely directed edges each having capacity C_{uv}. More generally, the capacity region has the form shown in Fig. 1.

Equation (5) characterizes the DMN *channel*. A DMN has more elements and rules associated with it, and we proceed to list these next (see [**K03**, § III.A-B]).

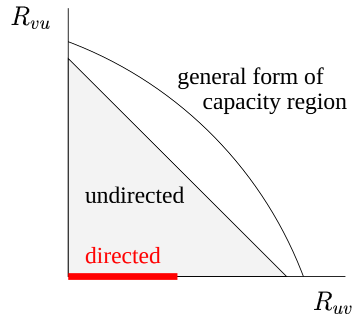

FIGURE 1. Capacity regions of different types of edges.

- The network is *clocked*, i.e., a universal clock ticks N times. Vertex u can transmit a symbol $X_{uv}^{(n)}$ into its TWC (u, v) *after* clock tick $n-1$ and *before* clock tick n for $n = 1, 2, \ldots, N$. An output symbol $Y_{vu}^{(n)}$ is received *at* clock tick n, i.e., there is a small delay in receiving symbols that ensures the network operates in a *causal* fashion.
- There are M independent messages $W^{(m)}$, $m = 1, 2, \ldots, M$, in the network. Message $W^{(m)}$ has NR_m bits so its *rate* is R_m bits per clock tick. Each message originates at exactly one vertex, but can be destined for any of the other $T-1$ vertices. Thus, each vertex has up to $2^{T-1}-1$ messages, one for each of the of the $2^{T-1}-1$ non-empty subsets of the other $T-1$ vertices.
- Let W_u be the set of messages originating at vertex u. The input $X_{uv}^{(n)}$ is a function of W_u and vertex u's past channel outputs

$$Y_u^{n-1} = Y_u^{(1)}, Y_u^{(2)}, \ldots, Y_u^{(n-1)}. \tag{6}$$

  Note that $Y_u^{(n)}$ is a vector that includes the nth channel outputs from *all* edges incident to u. Note also that $X_{uv}^{(n)}$ is *any* function of W_u and Y_u^{n-1}, so that we are permitting *joint* channel coding, routing, and/or network coding. We distinguish between routing and network coding in that routing permits message symbols and arriving *packets* (groups of input or output symbols) to be stored, re-ordered, and collected into other packets. Network coding, however, additionally allows packets to be *combined* to create new packets.
- Consider networks of TWCs. The channel outputs $Y_{uv}^{(n)}$ and $Y_{vu}^{(n)}$ are noisy functions of the channel inputs $X_{uv}^{(n)}$ and $X_{vu}^{(n)}$, i.e., we have

$$Y_{uv}^{(n)} = f_{uv}(X_{uv}^{(n)}, X_{vu}^{(n)}, Z_{uv}^{(n)}) \tag{7}$$
$$Y_{vu}^{(n)} = f_{vu}(X_{uv}^{(n)}, X_{vu}^{(n)}, Z_{uv}^{(n)}) \tag{8}$$

  for some functions $f_{uv}(\cdot)$ and $f_{vu}(\cdot)$, where $Z_{uv}^{(n)}$ is a noise random variable that is statistically independent of all other random variables. The functional relations (7) and (8) imply that (4) reduces to (5).
- Each message is decoded by one or more terminals. Suppose $W^{(m)}$ is destined for vertex v and is decoded as $\hat{W}_v^{(m)}$. Message estimate $\hat{W}_v^{(m)}$ is a function of vertex v's messages W_v and its channel outputs Y_v^N.

- A rate-tuple $(R_1, R_2, \ldots, R_M)$ is said to be *achievable* if there exist encoders and decoders such that

$$\Pr\left(\bigcup_{v,m}\left\{\hat{W}_v^{(m)} \neq W^{(m)}\right\}\right) < \epsilon$$

  for any positive ϵ. The *capacity region* $\mathcal{C}$ is the closure of the set of achievable rate-tuples.

3. A Flow Cut-set Bound

The celebrated max-flow/min-cut theorem of network optimization provides a useful outer bound on the set of feasible flow rates [**FF56**]. This bound partitions the vertices into two sets $\mathcal{S}$ and $\mathcal{S}^C$. Let $\mathcal{E}$ be a *cut*, i.e., a set of edges that separates the vertices in $\mathcal{S}$ from the vertices in $\mathcal{S}^C$. We further define $R_{\mathcal{S}\to\mathcal{S}^C}$ to be the sum of the rates of the messages originating at vertices in $\mathcal{S}$ and destined for one or more vertices in $\mathcal{S}^C$.

Suppose for the moment that $\mathcal{E}$ has only directed or undirected edges, and let $C_{uv} = 0$ if there is no edge between u and v, or if edge (u, v) is directed from v to u. The flow cut-set bound can be stated as follows (see [**S03**, Eq. (70.20) on p. 1228]).

PROPOSITION 1 (Flow Cut-set Bound). *For a network of directed and/or undirected edges, the feasible flow rates $(R_1, R_2, \ldots, R_M)$ satisfy, for all $\mathcal{S} \subseteq \mathcal{V}$,*

$$R_{\mathcal{S}\to\mathcal{S}^C} \leq \sum_{u\in\mathcal{S}, v\in\mathcal{S}^C} C_{uv} \tag{9}$$

$$R_{\mathcal{S}\to\mathcal{S}^C} + R_{\mathcal{S}^C\to\mathcal{S}} \leq \sum_{u\in\mathcal{S}, v\in\mathcal{S}^C} \max(C_{uv}, C_{vu}). \tag{10}$$

We remark that a bound such as (9) on $R_{\mathcal{S}^C\to\mathcal{S}}$ will appear when the set $\mathcal{S}$ is replaced with $\mathcal{S}^C$. We also remark that the flow cut-set bound generalizes to TWC edges by making C_{uv} the maximum possible rate from u to v.

4. A Bidirected Cut-set Bound

The *flow cut-set* bound is based on a "flow conservation" law, and hence may not apply to transmissions using network coding or transmissions across noisy channels. Instead, we must employ an *information-theoretic cut-set* bound that takes channel and network coding into account.

Our approach is to optimize the standard bound of [**CT91**, § 14.10]. The main object of the optimization is the following outer bound on the capacity region of a *single* TWC. We have $\mathcal{C}_{uv} \subseteq \overline{\mathcal{C}}_{uv}$, where $\overline{\mathcal{C}}_{uv}$ is the set of (R_{uv}, R_{vu}) satisfying

$$0 \leq R_{uv} \leq I(X_{uv}\,;\,Y_{uv} \mid X_{vu}) \tag{11}$$

$$0 \leq R_{vu} \leq I(X_{vu}\,;\,Y_{vu} \mid X_{uv}) \tag{12}$$

for any $P(x_{uv}, x_{vu})$. The set $\overline{\mathcal{C}}_{uv}$ is simply *Shannon's outer bound* on $\mathcal{C}_{uv}$. We remark that this bound can be loose, as was shown for binary multiplying channels in [**ZBS86, HW89**]. Fortunately, for several widely studied TWCs we do have $\overline{\mathcal{C}}_{uv} = \mathcal{C}_{uv}$. For example, this is true for noisy directed channels, push-to-talk TWCs, and Gaussian TWCs (see also [**S61**, § 11]).

We return to our bound. Summarizing the result, the bound reduces to the following three steps.

1. Convert every edge (u, v) into a pair of oppositely directed edges whose capacity pair is a boundary point of $\overline{\mathcal{C}}_{uv}$.
2. Apply the flow cut-set bound to get a rate region $\mathcal{R}'_{cut}$.
3. Repeat the above two steps for all boundary points of $\overline{\mathcal{C}}_{uv}$ on all edges. The union of the $\mathcal{R}'_{cut}$ is an outer bound $\mathcal{R}_{cut}$ on the capacity region $\mathcal{C}$.

Note that for directed networks, the above bound is the same as the flow cut-set bound and reduces to the results of [**B02**]. However, as shown below, the bounds can differ when there are two-way edges, e.g., undirected edges having triangular capacity regions. We call this information-theoretic bound a *bidirected* cut-set bound.

5. Implications of the Bidirected Bound

In what follows, when we refer to a *problem* we mean determining $\mathcal{C}$ when the messages originate at prescribed source vertices and are destined for prescribed sink vertices. Furthermore, when we refer to an *undirected* edge, we mean that this edge is a push-to-talk TWC. There are several consequences of the bidirected cut-set bound. First, the bound implies both (9) and (10), as summarized by following theorem.

THEOREM 5.1 (Region Ordering). *Consider a network of directed and/or undirected edges. Let $\mathcal{R}_{flow}$ be the feasible routing (flow) rates, $\mathcal{R}_{NC}$ the achievable network coding rates, $\mathcal{R}^{2D}_{cut}$ the bidirected cut-set rates, and $\mathcal{R}_{cut}$ the flow cut-set rates. We have the ordering*

$$\mathcal{R}_{flow} \subseteq \mathcal{R}_{NC} \subseteq \mathcal{R}^{2D}_{cut} \subseteq \mathcal{R}_{cut}. \tag{13}$$

PROOF. See (28) in the Appendix. □

The ordering (13) implies that any problem with $\mathcal{R}_{flow} = \mathcal{R}_{cut}$ also has $\mathcal{R}_{flow} = \mathcal{R}_{NC}$. Routing is therefore optimal in terms of rates. A survey of cases where $\mathcal{R}_{flow} = \mathcal{R}_{cut}$ can be found in [**S03**, Part VII] and a few such examples are discussed below.

Second, the bidirected bound involves converting a network of two-way channels into a set of *directed* networks. This observation leads to the following theorem.

THEOREM 5.2. *For any problem for which the flow cut-set bound for directed networks is achievable, one can also achieve the bidirected cut-set bound for networks of TWCs for which $\overline{\mathcal{C}}_{uv} = \mathcal{C}_{uv}$ for all (u, v).*

PROOF. By hypothesis, one achieves the flow cut-set for every directed network in the set of networks considered for the bidirected cut-set bound. But the bidirected bound is the union of the flow bounds for these directed networks, so any point in the bidirected bound can be achieved. □

For example, it is known that linear network coding is optimal for the *multicasting* problem in directed networks [**ACLY00, KM03**], where multicasting means a single message W is destined for a number of sinks. Linear network coding is therefore also optimal for multicasting in networks of TWCs for which $\overline{\mathcal{C}}_{uv} = \mathcal{C}_{uv}$ for all (u, v).

Third, observe that the bidirected cut-set bound treats each TWC *separately*, i.e., one needs to know only the rate-pairs of the individual capacity regions. This observation leads to the following result.

THEOREM 5.3. *For any problem for which the bidirected cut-set bound is achievable for a network of TWCs, one can separate channel and network coding.*

PROOF. Every point in the region defined by the bidirected cut-set bound corresponds to a point in the region of the flow cut-set bound for a noise-free directed network. We can mimic the operation of this network by using customized error-control codes over each TWC that drive the link error rates to zero. We then employ network coding as if the links were error-free, perhaps with probabilistic network codes if randomization is beneficial [**ACLY00**]. □

For example, we can achieve the bidirected cut-set bound for multicasting by employing capacity-approaching channel codes for each TWC link, and linear network coding on the network level. Stated in another way, it is rate-optimal to separate the physical layer design and network layer design when multicasting in networks of TWCs for which $\overline{\mathcal{C}}_{uv} = \mathcal{C}_{uv}$ for all (u, v).

We remark that the bidirected cut-set bound is valid for general communication problems, i.e., general source–sink scenarios. For example, consider the standard *multicommodity flow* problem treated in network optimization theory [**AMO93**, Ch. 17]. This problem has M commodities (messages) that each have one source vertex and one sink vertex [**AMO93**, Ch. 17]. The bidirected cut-set bound applies, and can potentially give tighter bounds than were available before, e.g., tighter bounds than the flow cut-set bound.

Finally, as we will see below, the bidirected cut-set bound can be loose. However, for directed networks one can sometimes determine improved outer bounds based on sets of disconnected edges. This is shown in Example 8 below.

6. Examples

Example 1: Consider a network with directed and/or undirected edges, and with one source and sink. The max-flow min-cut theorem of Ford and Fulkerson [**FF56**], Dantzig and Fulkerson [**DF56**], and Elias, Feinstein and Shannon [**EFS56**] ensures that $\mathcal{R}_{flow} = \mathcal{R}_{cut}$. This means that routing is optimal in terms of rates. Theorem 5.2 shows that the same is true for networks of two-way channels. Theorem 5.3 shows that one can separate channel and network coding.

Example 2: Consider any undirected network with two commodities. Hu's two-commodity (two-message) flow theorem [**H63**] ensures that max-biflow is min-cut. This means that network coding cannot improve on routing.

Example 3: Consider any undirected network that is planar, and that can be drawn in the plane so that all sources and sinks are on the boundary of the infinite region. A theorem of Okamura and Seymour [**OS81**] ensures that $\mathcal{R}_{flow} = \mathcal{R}_{cut}$. This means that routing is optimal.

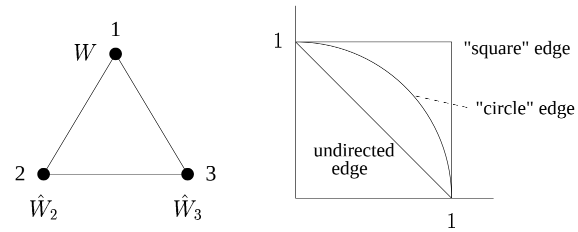

FIGURE 2. Multicast problem on different networks of TWCs.

Example 4: Consider the graph in Fig. 2 where the edges represent unit-capacity undirected edges. Suppose a message W of rate R is to be multicast from terminal 1 to terminals 2 and 3. The fbw cut-set bound is $R \leq 2$, and the bidirected cut-set bound gives

$$R \leq r_{12} + r_{13} \tag{14}$$

$$R \leq r_{12} + r_{32} \tag{15}$$

$$R \leq r_{13} + r_{23} \tag{16}$$

where $r_{ij} + r_{ji} \leq 1$ for $1 \leq i, j \leq 3$. The sum of (15) and (16) yields

$$2R \leq r_{12} + r_{13} + (r_{23} + r_{32}) \leq 3. \tag{17}$$

Thus, the bidirected cut-set bound gives $R \leq 3/2$, which is tighter than the fbw cut-set bound. $R = 3/2$ is achievable with routing as follows. One splits the message W into three parts $W^{(1)}$, $W^{(2)}$, $W^{(3)}$ that each have rate $1/2$. $W^{(1)}$ and $W^{(2)}$ are sent to terminal 2, and $W^{(1)}$ and $W^{(3)}$ are sent to terminal 3. Finally, $W^{(2)}$ and $W^{(3)}$ are exchanged on the TWC between terminals 2 and 3.

Example 5: Consider again the multicasting problem in Fig. 2, but suppose the capacity regions are the set of (R_{uv}, R_{vu}) satisfying

$$R_{uv}^2 + R_{vu}^2 \leq 1. \tag{18}$$

Suppose further that $\overline{\mathcal{C}}_{uv} = \mathcal{C}_{uv}$ for all (u, v). The boundary of this region is labeled "circle" edge in Fig. 2. The fbw cut-set bound is $R \leq 2$, and the bidirected cut-set bound gives (14)–(16) where $r_{ij}^2 + r_{ji}^2 \leq 1$ for $1 \leq i, j \leq 3$. The sum of (15) and (16) gives

$$2R \leq r_{12} + r_{13} + (r_{23} + r_{32}) \leq 2 + \sqrt{2}. \tag{19}$$

The rate $R = 1 + 1/\sqrt{2}$ is achievable by using the same routing strategy as in Example 4 but give $W^{(1)}$ rate $1 - 1/\sqrt{2}$, and $W^{(2)}$ and $W^{(3)}$ rate $1/\sqrt{2}$.

Example 6: Consider yet again the multicasting problem in Fig. 2, but suppose the capacity regions are the set of (R_{uv}, R_{vu}) satisfying

$$0 \leq R_{uv} \leq 1 \tag{20}$$

$$0 \leq R_{vu} \leq 1. \tag{21}$$

The boundary of this region is labeled "square" edge in Fig. 2. Both the fbw and bidirected cut-set bounds are $R \leq 2$, and $R = 2$ is clearly achievable with routing.

Example 7: We generalize Example 4 to a ring with T vertices where every edge is undirected with unit capacity. Suppose that vertex 1 multicasts the message W of rate R to K other vertices.

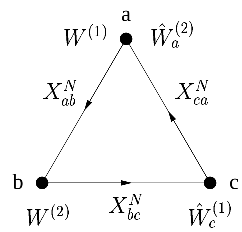

FIGURE 3. A directed triangular network.

Consider the following routing protocol. The source splits the message into $K+1$ parts that each have rate $1/K$. Label these parts $W^{(1)}, W^{(2)}, \ldots, W^{(K+1)}$. The source transmits $W^{(1)}, W^{(2)}, \ldots, W^{(K)}$ in the clockwise direction, and $W^{(2)}, W^{(3)}, \ldots, W^{(K+1)}$ in the counterclockwise direction. Each sink strips off the $W^{(i)}$ with the *largest* index i from the set of $W^{(i)}$ coming from the clockwise direction, and sends the remaining parts on in the same direction. Similarly, each sink strips off the $W^{(i)}$ with the *smallest* index i from the set of $W^{(i)}$ coming from the counterclockwise direction, and sends the remaining parts on in the same direction. The $T-(K+1)$ vertices that are not sources or sinks simply pass on what they receive. One can verify that all sinks receive all $K+1$ parts without violating the edge capacity constraints. The multicasting rate is therefore $R=(K+1)/K$.

For the bidirected cut-set bound, we successively group the $T-(K+1)$ vertices that are not sources or sinks with one of their neighboring vertices. We view the vertices in the same group as one vertex, and we view the original graph as a ring having $K+1$ vertices (the reason for doing this is so that we consider the same $K+1$ edges in every step of what follows). Consider those K cuts that have exactly one sink on one side of the cut, and the other K vertices on the other side of the cut. This gives the K rate bounds

$$\begin{aligned} R &\le r_{(i-1)i} + r_{(i+1)i}, \quad i = 2, 3, \ldots, K \\ R &\le r_{K(K+1)} + r_{1(K+1)} \end{aligned}$$

if the (grouped) vertices are numbered sequentially as one moves clockwise around the ring, vertex 1 is the source, and r_{ij} is the transmission rate from vertex i to vertex j. Summing these K bounds gives

$$\begin{aligned} KR &\le r_{12} + (K-1) + r_{1(K+1)} \\ &\le K+1 \end{aligned}$$

and so routing is rate-optimal for this problem.

Example 8: Consider a multicommodity flow problem on the network in Fig. 3 that has unit-capacity directed edges, and where $Y_{uv}^{(n)} = X_{uv}^{(n)}$ for all (u,v) and n. Suppose there are two messages: $W^{(1)}$ at vertex a destined for vertex c, and $W^{(2)}$ at vertex b destined for vertex a. The maximum sum rate is one because edge (b,c) must be used by both commodities. However, the bidirected cut-set bound merely constrains each commodity to have a rate of at most one, and this does not rule out a sum rate of two.

To derive a better bound, we will use Fano's inequality [**CT91**, § 8.9] and other basic information theoretic arguments to prove that the best sum rate is indeed one. For reliable communication, we have

$$
\begin{aligned}
N(R_1+R_2) &\overset{(a)}{\leq} I(W^{(1)};\hat{W}^{(1)}) + I(W^{(2)};\hat{W}^{(2)}) \\
&\overset{(b)}{\leq} I(W^{(1)};X_{bc}^N) + I(W^{(2)};W^{(1)}\ X_{ca}^N) \\
&\overset{(c)}{=} I(W^{(1)};X_{bc}^N) + I(W^{(2)};X_{ca}^N \mid W^{(1)}) \\
&\overset{(d)}{\leq} I(W^{(1)};X_{bc}^N) + I(W^{(2)};X_{bc}^N \mid W^{(1)}) \\
&\overset{(e)}{=} I(W^{(1)}W^{(2)};X_{bc}^N) \\
&\overset{(f)}{=} H(X_{bc}^N) - H(X_{bc}^N \mid W^{(1)}W^{(2)}) \\
&\overset{(g)}{\leq} H(X_{bc}^N) \\
&\overset{(h)}{\leq} N. \qquad (22)
\end{aligned}
$$

where (a) follows by Fano's inequality, (b) because $\hat{W}^{(1)}$ is a function of X_{bc}^N, $\hat{W}^{(2)}$ is a function of $W^{(1)}$ and X_{ca}^N, and by the Data Processing Inequality [**CT91**, § 2.8], (c) by the independence of the messages, (d) since X_{ca}^N is a function of X_{bc}^N, and by the Data Processing Inequality, (e) by the chain rule for mutual information [**CT91**, Theorem 2.5.2], (f) by the definition of mutual information [**CT91**, p. 20], (g) by the non-negativity of entropy [**CT91**, Lemma 2.1.1], and (h) since the channel (b,c) has unit capacity. This bound is a kind of *disconnecting edge set bound*, i.e., a bound based a set of edges that disconnects sources from sinks. It is interesting to consider whether one can generalize the bound.

Acknowledgments

The authors wish to thank C. Chekuri and B. Shepherd of Bell Laboratories for useful discussions, and C. Fragouli of EPFL Switzerland for comments on the manuscript.

APPENDIX A
PROOF OF THE BIDIRECTED CUT-SET BOUND

We partition the vertices $\mathcal{V}$ into two sets $\mathcal{S}$ and $\mathcal{S}^C$. Let $Y_{\mathcal{S}}$ denote $\{Y_u : u \in \mathcal{S}\}$, let $Y_{\mathcal{S}\to\mathcal{T}}$ denote $\{Y_{uv} : u \in \mathcal{S}, v \in \mathcal{T}\}$, and similarly for $X_{\mathcal{S}}$ and $X_{\mathcal{S}\to\mathcal{T}}$. The cut-set bound in [**CT91**, § 14.10] states that, for reliable communication, we have

$$R_{\mathcal{S}\to\mathcal{S}^C} \leq I(X_{\mathcal{S}}\,;\,Y_{\mathcal{S}^C} \mid X_{\mathcal{S}^C}) \qquad (23)$$

for *all* $\mathcal{S} \in \mathcal{V}$ and for some joint probability distribution $P(x_{\mathcal{V}})$ for $X_{\mathcal{V}}$. Note that (23) involves only *one* use of the network, i.e., the $Y_{\mathcal{V}}$ are all defined through the conditional probability distribution (5). To state the bound more precisely, let $\mathcal{R}'(P(x_{\mathcal{V}}),\mathcal{S})$ be the set of (non-negative) rate-tuples $(R_1, R_2 \ldots, R_M)$ that are permitted by (23). The cut-set bound in [**CT91**, § 14.10] states that the capacity region $\mathcal{C}$ satisfies

$$\mathcal{C} \subseteq \bigcup_{P(x_{\mathcal{V}})} \bigcap_{\mathcal{S}\subseteq\mathcal{V}} \mathcal{R}'(P(x_{\mathcal{V}}),\mathcal{S}). \qquad (24)$$

We emphasize that (24) involves first an intersection of regions and then a union, and not the other way around.

We continue by manipulating (23). We have

$$
\begin{aligned}
R_{\mathcal{S}\to\mathcal{S}^C} &\overset{(a)}{\leq} I(X_{\mathcal{S}}\,;\,Y_{\mathcal{S}\to\mathcal{S}^C}Y_{\mathcal{S}^C\to\mathcal{S}^C}\mid X_{\mathcal{S}^C}) \\
&\overset{(b)}{=} I(X_{\mathcal{S}}\,;\,Y_{\mathcal{S}\to\mathcal{S}^C}\mid X_{\mathcal{S}^C}) \\
&\overset{(c)}{=} H(Y_{\mathcal{S}\to\mathcal{S}^C}\mid X_{\mathcal{S}^C}) - H(Y_{\mathcal{S}\to\mathcal{S}^C}\mid X_{\mathcal{V}}) \\
&\overset{(d)}{=} H(Y_{\mathcal{S}\to\mathcal{S}^C}\mid X_{\mathcal{S}^C}) - \sum_{u\in\mathcal{S},v\in\mathcal{S}^C} H(Y_{uv}\mid X_{uv}X_{vu}) \\
&\overset{(e)}{\leq} \sum_{u\in\mathcal{S},v\in\mathcal{S}^C} H(Y_{uv}\mid X_{vu}) - H(Y_{uv}\mid X_{uv}X_{vu}) \\
&\overset{(f)}{=} \sum_{u\in\mathcal{S},v\in\mathcal{S}^C} I(X_{uv}\,;\,Y_{uv}\mid X_{vu}) \qquad (25)
\end{aligned}
$$

where (a) follows by (23) and by definition of $Y_{\mathcal{S}^C}$, (b) by Markovity (cf. (5)), (c) by definition of mutual information, (d) by Markovity, (e) by expanding the first entropy using the chain rule for mutual information, and by the fact that conditioning cannot increase entropy, and (f) by definition. We remark that equality holds in step (e) if the input on any edge is statistically independent of the input on any *other* edge. Furthermore, one can transform $P(x_{\mathcal{V}})$ into a new distribution

$$P'(x_{\mathcal{V}}) = \prod_{(u,v)} P(x_{uv}, x_{vu}) \qquad (26)$$

for which we have

$$\mathcal{R}'(P(x_{\mathcal{V}}), \mathcal{S}) \subseteq \mathcal{R}'(P'(x_{\mathcal{V}}), \mathcal{S}). \qquad (27)$$

Thus, one may as well restrict attention to those $P(x_{\mathcal{V}})$ whose inputs across the TWCs are statistically independent. The bound (24) thus simplifies to the bidirected bound described in Section 4.

For example, suppose we have a network of directed and/or undirected edges. The points on the boundary of $\mathcal{C}_{uv}$ of an undirected edge (u,v) can be written as $(\tau\, C_{uv}, (1-\tau)\, C_{uv})$ where $0 \leq \tau \leq 1$. We further write $\tau_{vu} = 0$ if an edge is directed from u to v. The bidirected cut-set bound is then

$$R_{\mathcal{S}\to\mathcal{S}^C} \leq \sum_{u\in\mathcal{S},v\in\mathcal{S}^C} \tau_{uv} C_{uv} \qquad (28)$$

for all $\mathcal{S} \subseteq \mathcal{V}$, and for some splitting of the undirected edges' rates, i.e., for each undirected edge we choose τ_{uv} and τ_{vu} so that $0 \leq \tau_{uv} \leq 1$ and $\tau_{uv} + \tau_{vu} = 1$. Inequality (28) clearly implies (9) and (10), as claimed at the beginning of Section 5.

References

[ACLY00] R. Ahlswede, N. Cai, S.-Y .R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1204–1216, July 2000.

[AMO93] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Englewood Cliffs: Prentice Hall, 1993.

[B02] S. P. Borade, "Network information flow: limits and achievability," *Proc. 2002 IEEE Int. Symp. Inform. Theory*, (Lausanne, Switzerland), p. 139, June 30-July 5, 2002.

[C75] A. B. Carleial, *On the Capacity of Multiple-Terminal Communication Networks*. Ph.D. thesis, Stanford Univ., Stanford, CA, Aug. 1975.

[CT91] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.

[DF56] G. B. Dantzig and D. R. Fulkerson, "On the max-flow, min-cut theorem of networks," in "Linear Inequalities," *Ann. Math. Studies*, no. 38, pp. 215-221, Princeton, New Jersey, 1956.

[EFS56] P. Elias, A. Feinstein, and C. E. Shannon, "A note on the maximum flow through a network," *IRE Trans. Inform. Theory*, vol. 2, pp. 117–119, Dec. 1956.

[FF56] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network," *Can. J. Math.*, vol. 8, pp. 399-404, 1956.

[HW89] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

[H63] T. C. Hu, "Multi-commodity network flows," *Oper. Res.*, vol. 11, pp. 344-360, 1963.

[KM03] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, 2003.

[K03] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inform. Theory*, vol. 49, no. 1 , pp. 4–21, Jan. 2003.

[vdM68] E. C. van der Meulen, *Transmission of Information in a T-Terminal Discrete Memoryless Channel*. Ph.D. thesis, Univ. of California, Berkeley, CA, June 1968.

[vdM77] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. IT–23, pp. 1–37, Jan. 1977.

[OS81] H. Okamura and P. D. Seymour, "Multicommodity flows in planar graphs," *J. Combin. Theory*, Series B 31, pp. 75–81, 1981.

[S03] A. Schrijver, *Combinatorial Optimization*, New York: Springer, 2003.

[S61] C.E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. on Mathematical Statistics and Probability*, J. Neyman, Ed., Berkeley, CA, 1961, vol. 1, pp. 611–644, Univ. Calif. Press, Reprinted in *Claude Elwood Shannon: Collected Papers*, pp. 351-384, (N.J.A. Sloane and A.D. Wyner, eds.) Piscataway: IEEE Press, 1993.

[ZBS86] Z. Zhang, T. Berger, and J. P. M. Schalkwijk, "New outer bounds to capacity regions of two-way channels," *IEEE Trans. Inform. Theory*, vol. 32, no. 3, pp. 383–386, May 1986.

MATHEMATICS OF COMMUNICATIONS RESEARCH DEPARTMENT, BELL LABORATORIES, LUCENT TECHNOLOGIES, 600 MOUNTAIN AVE., MURRAY HILL, NEW JERSEY 07974

E-mail address: `gkr@research.bell-labs.com`

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109

E-mail address: `savari@eecs.umich.edu`