An information-theoretic meta-theorem on edge-cut bounds

Sudeep Kamath
EECS Dept., UC Berkeley
Berkeley, CA 94720
Email: sudeep@eecs.berkeley.edu

Pramod Viswanath
CSL and Dept. of ECE
UIUC, IL 61801
Email: pramodv@illinois.edu

***Abstract*—We consider the problem of multiple unicast in wireline networks. Edge-cut based bounds which are simple bounds on the rates achievable by routing flow are not in general, fundamental, i.e. they are not outer bounds on the capacity region. It has been observed that when the problem has some kind of symmetry involved, then flows and edge-cut based bounds are 'close', i.e. within a constant or poly-logarithmic factor of each other. In this paper, we make the observation that in these very cases, such edge-cut based bounds are actually 'close' to fundamental yielding an approximate characterization of the capacity region for these problems. We demonstrate this in the case of k-unicast in undirected networks, k-pair unicast in directed networks with symmetric demands i.e. for every source communicating to a destination at a certain rate, the destination communicates an independent message back to the source at the same rate, and sum-rate of k-groupcast in directed networks, i.e. a group of nodes, each of which has an independent message for every other node in the group. We place our work in context of existing results to suggest a meta-theorem: if there is inherent symmetry either in the network connectivity or in the traffic pattern, then edge-cut bounds are near-fundamental and flows approximately achieve capacity.**

I. Introduction

The central problem of network information theory is to characterize the capacity region of a general network. Wireline networks are a special class of such networks where the edges between vertices are *orthogonal* and *noise-free*. In this class of networks, network coding has the potential to provide significant advantages in comparison to flow (i.e. routing strategies) for multicast problems [1] as well as for multiple unicast problems [2]. Recent results due to Dougherty, Freiling, Zeger and Chan, Grant suggest that characterizing the capacity region of a multiple unicast network is a hard problem [3], [4], [5]. In particular, even coding strategies such as linear codes do not achieve capacity in general [4].

On the other hand, the literature on hardness of cut problems typically deal with edge-cut bounds which are conventional outer bounds on flow. But these bounds are not fundamental bounds on the capacity region [6], i.e. they can often be beaten if network coding is allowed. Although edge-cut bounds in directed networks are not fundamental, they are combinatorially well-represented. They are however, hard to approximate in general [7], [8].

One class of networks for which edge-cut bounds can be approximated well are undirected networks. Leighton, Rao [9] and Linial, London, Rabinovich [10] show that for the problem of k-unicast in undirected networks, flow solutions approach the edge-cut bounds up to a factor of $\Theta(\log k)$. There has also been discovered a semi-definite programming relaxation approach that allows an approximation of edge-cut bounds up to a factor of $\Theta(\sqrt{\log k}\log\log k)$ [11]. Interestingly, for undirected networks, edge-cut bounds can be derived from the vertex bipartition cutset bound and are hence, fundamental outer bounds on the capacity region. Thus, [9], [10] also characterize up to a factor of $\Theta(\log k)$ the capacity region of k-unicast in undirected networks. It has been conjectured that flow solutions in fact, achieve capacity.

Another setting in which edge-cut bounds can be approximated well is the problem of multiple unicast in directed wireline networks with *symmetric demands*, i.e. for each source communicating to its destination at a certain rate, there is an independent message to be communicated from the destination back to the source at the same rate. Klein, Plotkin, Rao, Tardos [14] show under this model that flow solutions achieve within $\Theta(\log^2 k)$ of the edge-cut bounds. We ask the question: "Are these edge-cut bounds fundamental outer bounds on the capacity region?" Surprisingly, the answer turns out to be yes and the proof of this result is one of the main contributions of this paper. This completes an approximate characterization of the capacity region for this class of problems. The key tool we use in the proof is the Generalized Network Sharing (GNS) bound that was first developed in [15] for directed wireline networks and was also used subsequently for two-unicast linear deterministic networks [16].

Another interesting setting is that of groupcast in directed wireline networks. There is a group of nodes and each node in the group has one independent message for each other node in the group. Naor, Zosin [17] show that the maximum sum-rate achievable by routing flow for groupcast is at least half the *multicut*, a simple edge-cut based outer bound on flow. We ask the question: "Is the multicut a fundamental upper bound on the sum-rate?". We find that the answer is no, but that twice the multicut is indeed a fundamental upper bound. This shows that routing flow approximately achieves the sum-rate-capacity for groupcast networks.

When there is some kind of symmetry in the network, either in the underlying graph (undirected or bidirected networks) or in the traffic (directed network with symmetric demands or groupcast sum-rate), the following picture seems to emerge.

- (Achievability) *Algorithmic Meta-Theorem*: Edge-cut bounds can be well-approximated either by flows [9], [10], [14], [17], [18] or by other means [11].
- (Converse) *Information-Theoretic Meta-Theorem*: Edge-cut bounds are near-fundamental outer bounds on the capacity region.
- *Combined Meta-Theorem*: Flow approximately achieves capacity.

In a companion paper [19], we use the results of this paper and achievability results similar to [14] which have been obtained for the class of polymatroidal networks [18] to study the capacity regions of multiple unicast with symmetric demands in different classes of Gaussian networks. The rest of this paper is organized as follows. We set up notation and preliminaries in Section II. We briefly state the Generalized Network Sharing (GNS) bound in Section III. We then discuss k-unicast undirected networks in Section IV, k-pair unicast directed symmetric-demand networks in Section V and k-groupcast directed networks in Section VI. We finally conclude with a discussion in Section VII.

II. Preliminaries

Definition. A k-unicast directed network $\mathcal{N}$ for source-destination pairs $\{(s_i;d_i)\}_{i\in\mathcal{I}}$ with $|\mathcal{I}|=k$ (for instance, $\mathcal{I}:=\{1,2,\ldots,k\}$) is a tuple $(\mathcal{G},\underline{\mathrm{C}})$ where

- $\mathcal{G}=(\mathcal{V},\mathcal{E})$ is the underlying directed graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$, with $s_i,d_i\in\mathcal{V}(\mathcal{G})$ for $i\in\mathcal{I}$,
- $\underline{\mathrm{C}}=(C_e : e\in\mathcal{E}(\mathcal{G}))$ is the edge-capacity vector, with $C_e\in\mathbb{R}_{\geq 0}\cup\{\infty\}\ \forall e\in\mathcal{E}(\mathcal{G})$.

For each $i\in\mathcal{I}$, s_i has independent information to be communicated to d_i at rate R_i.

Notation. For $v\in\mathcal{V}(\mathcal{G})$, let $\mathsf{In}(v)$ and $\mathsf{Out}(v)$ denote the edges entering into and leaving v respectively.

Definition. Given a k-unicast directed network $\mathcal{N}=(\mathcal{G},\underline{\mathrm{C}})$ for source-destination pairs $\{(s_i;d_i)\}_{i\in\mathcal{I}}$, we say that the non-negative rate tuple $(R_i : i\in\mathcal{I})$ is *achievable* if for any $\epsilon>0$, there exist positive integers N and T (called block length and number of epochs respectively), a finite alphabet $\mathcal{A}$ with $|\mathcal{A}|\geq 2$ and using notation $H_v:=\Pi_{i\in\mathcal{I}:v=s_i}\mathcal{A}^{\lceil NTR_i\rceil}$ (with an empty product being the singleton set),

- encoding functions for $1\leq t\leq T, e=(u,v)\in\mathcal{E}$, $f_{e,t}: H_u\times\Pi_{e'\in\mathsf{In}(u)}\left(\mathcal{A}^{\lfloor NC_{e'}\rfloor}\right)^{(t-1)}\mapsto\mathcal{A}^{\lfloor NC_e\rfloor}$,
- decoding functions at destinations d_i for $i\in\mathcal{I}$, $f_{d_i}: H_{d_i}\times\Pi_{e'\in\mathsf{In}(d_i)}\left(\mathcal{A}^{\lfloor NC_{e'}\rfloor}\right)^{T}\mapsto\mathcal{A}^{\lceil NTR_i\rceil}$,

with the property that under the uniform probability distribution on $\Pi_{i\in\mathcal{I}}\mathcal{A}^{\lceil NTR_i\rceil}$,

$$\mathsf{Pr}\left(g(m_1,m_2,\ldots,m_k)\neq(m_1,m_2,\ldots,m_k)\right)\leq\epsilon,$$

where $g:\Pi_{i\in\mathcal{I}}\mathcal{A}^{\lceil NTR_i\rceil}\mapsto\Pi_{i\in\mathcal{I}}\mathcal{A}^{\lceil NTR_i\rceil}$ is the global decoding function induced inductively by $\{f_{e,t}: e\in\mathcal{E}(\mathcal{G}), 1\leq t\leq T\}$ and $\{f_{d_i} : i\in\mathcal{I}\}$. The closure of the set of achievable rate tuples is called the *capacity region* and is denoted by $\mathcal{C}$. Define the *sum-rate-capacity* by $C_{\text{sum}-\text{rate}}:=\sup_{(R_i:i\in\mathcal{I})\in\mathcal{C}}\sum_{i\in\mathcal{I}}R_i$.

Definition. Given a k-unicast directed network $\mathcal{N}=(\mathcal{G},\underline{\mathrm{C}})$ for source-destination pairs $\{(s_i;d_i)\}_{i\in\mathcal{I}}$, we say that the non-negative rate tuple $(R_i : i\in\mathcal{I})$ is *achievable by routing flow* if there exist for each $i\in\mathcal{I}$ and each $e=(u,v)\in\mathcal{E}(\mathcal{G})$, real numbers $f_{i,e}\geq 0$ such that $\sum_{i\in\mathcal{I}}f_{i,e}\leq C_e\ \forall\ e\in\mathcal{E}(\mathcal{G})$, and for each $i\in\mathcal{I}$ and each $v\in\mathcal{V}(\mathcal{G})$,

$$\sum_{e\in\mathsf{Out}(v)}f_{i,e}-\sum_{e\in\mathsf{In}(v)}f_{i,e}=\begin{cases}0 & \text{if } v\neq s_i,d_i,\\ R_i & \text{if } v=s_i,\\ -R_i & \text{if } v=d_i.\end{cases}$$

The closure of the set of rate tuples achievable by routing flow is called the *flow region* and is denoted by $\mathcal{F}$. Define the *sum-rate-max-flow* by $F_{\text{sum}-\text{rate}}:=\sup_{(R_i:i\in\mathcal{I})\in\mathcal{F}}\sum_{i\in\mathcal{I}}R_i$

Definition. Given a k unicast directed network $\mathcal{N}=(\mathcal{G},\underline{\mathrm{C}})$ for source-destination pairs $\{(s_i;d_i)\}_{i\in\mathcal{I}}$, we define the *edge-cut outer bound*, denoted by $\mathcal{R}_{\text{edge}-\text{cut}}$, to be the set of all non-negative tuples $(R_i : i\in\mathcal{I})$ that satisfy for every $E\subseteq\mathcal{E}(\mathcal{G})$, the inequality $\sum_{i\in J}R_i\leq\sum_{e\in E}C_e$ where index $i\in J\subseteq\mathcal{I}$ if and only if $\mathcal{G}\setminus E$ has no directed paths from s_i to d_i. We define the *multicut*, denoted by R_{multicut}, to be the minimum value of $\sum_{e\in E}C_e$ over all $E\subseteq\mathcal{E}(\mathcal{G})$ with the property that $\mathcal{G}\setminus E$ has no directed paths from s_i to d_i for each $i\in\mathcal{I}$.

Remark 1. R_{multicut} may in general be strictly larger than the tighter bound on $F_{\text{sum}-\text{rate}}$ given by $\sup_{(R_i:i\in\mathcal{I})\in\mathcal{R}_{\text{edge}-\text{cut}}}\sum_{i\in\mathcal{I}}R_i$.

While it is clear that $\mathcal{F}\subseteq\mathcal{R}_{\text{edge}-\text{cut}}$ and $\mathcal{F}\subseteq\mathcal{C}$, the connection between $\mathcal{C}$ and $\mathcal{R}_{\text{edge}-\text{cut}}$ is unclear. It is easy to show examples where $\mathcal{C}\nsubseteq\mathcal{R}_{\text{edge}-\text{cut}}$. Thus, simple edge-cut based outer bounds are not in general, *fundamental*, i.e. they are not outer bounds on the capacity region. Likewise it is clear that $F_{\text{sum}-\text{rate}}\leq R_{\text{multicut}}$ and $F_{\text{sum}-\text{rate}}\leq C_{\text{sum}-\text{rate}}$ but $C_{\text{sum}-\text{rate}}$ and R_{multicut} have no apparent connection. Indeed, [2] provides a series of k-unicast networks, one for each k with $k=2^n$ with $F_{\text{sum}-\text{rate}}=R_{\text{multicut}}=\frac{1}{k}C_{\text{sum}-\text{rate}}$ and $\mathcal{C}\nsubseteq(k-\epsilon)\mathcal{R}_{\text{edge}-\text{cut}}$ for any $\epsilon>0$. [20] shows that the gap between $\mathcal{F}$ and $\mathcal{R}_{\text{edge}-\text{cut}}$ can be as large as $k-\epsilon$ for any $\epsilon>0$.

III. Generalized Network Sharing (GNS) bound

Definition. Given a directed network $\mathcal{N}=(\mathcal{G},\underline{\mathrm{C}})$ with a set of $2r$ distinguished vertices $w_1,w_2,\ldots,w_r,w'_1,w'_2,\ldots,w'_r$, if a set of edges $E\subseteq\mathcal{E}(\mathcal{G})$ has the property that $\mathcal{G}\setminus E$ has no directed paths from w_i to w'_j whenever $\pi(i)\geq\pi(j)$, $1\leq i,j\leq r$, for some permutation $\pi:\{1,2,\ldots,r\}\mapsto\{1,2,\ldots,r\}$, then we say that E is a *GNS-cut* for $\{w_1,w_2,\ldots,w_r;w'_1,w'_2,\ldots,w'_r\}$.

Theorem 1. *(Generalized Network Sharing (GNS) bound from Kamath-Tse-Anantharam [15])*

Let $\mathcal{N}=(\mathcal{G},\underline{\mathrm{C}})$ be a k-unicast directed network for source-destination pairs $\{s_i;d_i\}_{i\in\mathcal{I}}$. If, for $J=\{j_1,j_2,\ldots,j_r\}\subseteq\mathcal{I}$ and $E\subseteq\mathcal{E}(\mathcal{G})$, we have that E is a GNS-cut for

$\{s_{j_1}, s_{j_2}, \ldots, s_{j_r}; d_{j_1}, d_{j_2}, \ldots, d_{j_r}\}$, *then for all* $(R_i : i \in \mathcal{I}) \in \mathcal{C}$, *we have* $\sum_{j \in J} R_j \leq \sum_{e \in E} C_e$.

We skip the proof of Theorem 1 due to lack of space. The essential idea is contained in [15]. The GNS bound is to the capacity region what the edge-cut bound is to the commodity flow region, namely an intuitive outer bound that arises from simple connectivity properties of the underlying graph of the network.

IV. k-UNICAST UNDIRECTED NETWORKS

We have skipped the natural counterpart definitions of $\mathcal{F}, \mathcal{C}, \mathcal{R}_{\text{edge-cut}}$ for undirected networks due to lack of space. All statements in this section refer to k-unicast undirected networks.

Theorem 2. *(Leighton-Rao [9], Linial-London-Rabinovich [10])*

$$\frac{\mathcal{R}_{\text{edge-cut}}}{\Theta(\log k)} \subseteq \mathcal{F} \subseteq \mathcal{R}_{\text{edge-cut}}. \tag{1}$$

Theorem 3.

$$\mathcal{C} \subseteq \mathcal{R}_{\text{edge-cut}}. \tag{2}$$

Theorem 3 follows from the vertex bipartition cutset bound and we omit the proof here. Theorems 2 and 3 together imply that routing flow is approximately capacity-achieving:

Corollary 4.

$$\frac{\mathcal{R}_{\text{edge-cut}}}{\Theta(\log k)} \subseteq \mathcal{F} \subseteq \mathcal{C} \subseteq \mathcal{R}_{\text{edge-cut}}. \tag{3}$$

Indeed, the *Li and Li conjecture* states that flow achieves capacity for k-unicast in undirected graphs.

Conjecture 5. *(Li-Li [12], Harvey-Kleinberg-Lehman [13])*

$$\mathcal{F} = \mathcal{C}.$$

V. k-PAIR UNICAST DIRECTED SYMMETRIC-DEMAND NETWORKS

Definition. A *k-pair unicast directed symmetric-demand network* is a $2k$-unicast directed network $\mathcal{N}$ with $2k$ distinct distinguished nodes (source-destination nodes) $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$ with source-destination pairs $\{s_i; d_i\}_{i \in \mathcal{I}}$ where $\mathcal{I} = \{1, 2, \ldots, k\} \cup \{-1, -2, \ldots, -k\}$ and for $i > 0$, $s_i = u_i$, $d_i = v_i$, while for $i < 0, s_i = v_{-i}$, $d_i = u_{-i}$. The rate tuple $(R_i : 1 \leq i \leq k)$ is defined to be in the capacity region $\mathcal{C}$, flow region $\mathcal{F}$, edge-cut outer bound $\mathcal{R}_{\text{edge-cut}}$ for the k-pair unicast directed symmetric-demand network if the rate tuple $(R'_i : i \in \mathcal{I})$, given by $R'_i = R_{|i|}$ for $i \in \mathcal{I}$, lies in the capacity region, flow region, edge-cut outer bound respectively of the $2k$-unicast directed network.

Remark 2. There is no loss of generality in assuming $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$ distinct since if they aren't, we can add more nodes and infinite capacity edges to make them distinct while obtaining a network with identical capacity region.

Definition. Given a k-pair unicast directed symmetric-demand network $\mathcal{N} = (\mathcal{G}, \underline{\mathrm{C}})$ with source-destination nodes $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$, we define the *GNS-cut outer bound*, denoted by $\mathcal{R}_{\text{GNS-cut}}$, to be the set of all non-negative tuples $(R_i : 1 \leq i \leq k)$ that satisfy for every $E \subseteq \mathcal{E}(\mathcal{G})$, the inequality $\sum_{i \in J} R_i \leq \sum_{e \in E} C_e$ whenever E is a GNS-cut for $\{w_1, w_2, \ldots, w_r; w'_1, w'_2, \ldots, w'_r\}$ where

- $J \subseteq \{1, 2, \ldots, k\}, |J| = r$,
- $w_1, w_2, \ldots, w_r, w'_1, w'_2, \ldots, w'_r$ are distinct,
- for $1 \leq j \leq r$, $(w_j, w'_j) = (u_i, v_i)$ or (v_i, u_i) for some $i \in J$.

We define a *weak edge-cut* outer bound for this class of networks.

Definition. Given a k-pair unicast directed symmetric-demand network $\mathcal{N} = (\mathcal{G}, \underline{\mathrm{C}})$ with source-destination nodes $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$, we define the *weak-edge-cut outer bound*, denoted by $\mathcal{R}_{\text{weak-edge-cut}}$, to be the set of all non-negative tuples $(R_i : 1 \leq i \leq k)$ that satisfy for every $E \subseteq \mathcal{E}(\mathcal{G})$, the inequality $\sum_{i \in J} R_i \leq \sum_{e \in E} C_e$ where index $i \in J \subseteq \{1, 2, \ldots, k\}$ if and only if $\mathcal{G} \setminus E$ has no directed paths from either u_i to v_i or v_i to u_i or both.

All statements in this section refer to k-pair unicast directed symmetric-demand networks.

Remark 3. It is easy to see that

$$\mathcal{R}_{\text{edge-cut}} \subseteq \mathcal{R}_{\text{weak-edge-cut}} \subseteq 2\,\mathcal{R}_{\text{edge-cut}}.$$

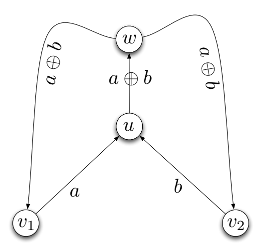


Fig. 1. All edges have unit capacity

Theorem 6. *(Klein-Plotkin-Rao-Tardos [14])*

$$\frac{\mathcal{R}_{\text{weak-edge-cut}}}{\Theta(\log^2 k)} \subseteq \mathcal{F} \subseteq \mathcal{R}_{\text{weak-edge-cut}}. \tag{4}$$

Theorem 7. *(follows from the GNS bound of Theorem 1)*

$$\mathcal{C} \subseteq \mathcal{R}_{\text{GNS-cut}}. \tag{5}$$

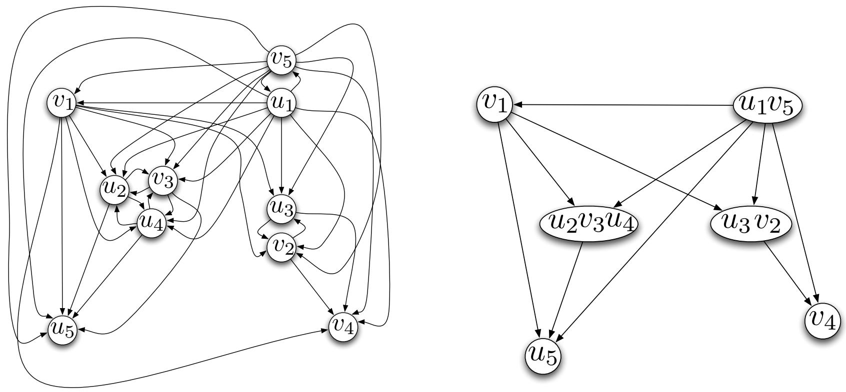
Theorem 8.

$$\mathcal{R}_{\text{weak-edge-cut}} = \mathcal{R}_{\text{GNS-cut}}. \tag{6}$$

Theorems 6, 7 and 8 together imply that routing flow is approximately capacity-achieving:

Corollary 9.

$$\frac{\mathcal{R}_{\text{weak-edge-cut}}}{\Theta(\log^2 k)} \subseteq \mathcal{F} \subseteq \mathcal{C} \subseteq \mathcal{R}_{\text{weak-edge-cut}} = \mathcal{R}_{\text{GNS-cut}}. \tag{7}$$



(a) Connectivity graph $\mathcal{G}_c$ (b) Reduced connectivity graph $\mathcal{G}_r$

Fig. 2. Connectivity graph and reduced connectivity graph for an instance of a 5-pair unicast directed symmetric-demand network

We conjecture that network coding can improve rates by at most a constant factor α for k-pair unicast symmetric demand networks, where α does not depend on k.

Conjecture 10. *(Analog of the Li and Li conjecture)*

$$\mathcal{F} \subseteq \mathcal{C} \subseteq \alpha\,\mathcal{F}.$$

For the 1-pair unicast symmetric demand network in Fig. 1 with $v_1 = s_1, v_2 = d_1$ the simple XOR coding scheme shows that if Conjecture 10 is true, then we must have $\alpha \geq 2$. It also shows that in general, $\mathcal{C} \not\subseteq \mathcal{R}_{\text{edge-cut}}$ for these networks.

Now, we prove the equivalence between weak-edge-cuts and GNS-cuts for k-pair unicast directed symmetric-demand networks, thus proving Theorem 8.

Proof: It is easy to see that the inequality obtained from a GNS-cut can always be obtained from a weak-edge-cut since a GNS-cut requires stronger disconnections as compared to a weak-edge-cut. This gives $\mathcal{R}_{\text{weak-edge-cut}} \subseteq \mathcal{R}_{\text{GNS-cut}}$. To show $\mathcal{R}_{\text{GNS-cut}} \subseteq \mathcal{R}_{\text{weak-edge-cut}}$, we now consider $E \subseteq \mathcal{E}(\mathcal{G})$, and say $i \in J \subseteq \{1, 2, \ldots, k\}$ if and only if $\mathcal{G} \setminus E$ has no directed paths from either u_i to v_i or from v_i to u_i or both. We show that E is a GNS-cut for $\{w_1, w_2, \ldots, w_r; w'_1, w'_2, \ldots, w'_r\}$ where the $2r$ vertices $w_1, w_2, \ldots, w_r, w'_1, w'_2, \ldots, w'_r$ are all distinct and for $1 \leq j \leq r$, $(w_j, w'_j) = (u_i, v_i)$ or (v_i, u_i) for some $i \in J$ with $|J| = r$. We will prove this for the case $J = \{1, 2, \ldots, k\}$. The proof for other choices of J is similar.

Define the *connectivity graph* $\mathcal{G}_c$ as a directed graph over $2k$ vertices $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$ as follows. For every pair of distinct vertices w and z, there is a directed edge from w to z in $\mathcal{G}_c$ if and only if w has a directed path to z in $\mathcal{G} \setminus E$. See Fig. 2(a) for an example. $\mathcal{G}_c$ is transitively closed, i.e. for three distinct vertices w, z, x, if w has an edge to z and z has an edge to x, then w has an edge to x.

Now, define the *reduced connectivity graph* $\mathcal{G}_r$ as the directed acyclic graph with vertices represented by the strongly connected components of $\mathcal{G}_c$. See Fig. 2(b) for an example. For each $i = 1, 2, \ldots, k$, we have that u_i and v_i do not lie in the same strongly connected component. Note that a directed acyclic graph has at least one sink vertex, i.e. a vertex with no outgoing edges. Consider the following algorithm that fills in the cells of an initially empty $k \times 2$ table with vertex names from $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$.

(1) Pick any sink vertex in directed acyclic graph $\mathcal{G}_r$.
(2) List the vertices of $\mathcal{G}_c$ in the strongly connected component represented by the chosen sink vertex.
   (a) Pick a vertex w from the list.
   (b) If vertex w has been entered previously in the table, do nothing. Else, add vertex w in the first column of the lowest row in the table not yet filled. Add the destination of vertex w in the second column of the same row, e.g. if v_3 was entered in the first column of the lowest available row, then fill u_3 in the second column.
   (c) Remove w from the list and go back to (a) if the list is still non-empty, else proceed to (3)
(3) Modify graph $\mathcal{G}_r$ by deleting the chosen sink vertex. If this graph has non-zero number of vertices, go to step (1), else quit.

It is easy to verify that the following properties hold upon termination of the algorithm. For an example, see Fig. 3.

(i) Each of $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k$ shows up exactly once in the table.
(ii) Each row of the table is made up of vertices u_i and v_i for some i.
(iii) In graph $\mathcal{G}_c$, vertex w obtained from the first column of row i does not have an edge to vertex z obtained from the second column of row j whenever $i \geq j$.

Now, if the j^{th} row of the table consists of u_i, v_i, we set $\pi(j) = i$ and $(w_i, w'_i) = (u_i, v_i)$ or (v_i, u_i) depending on whether the first entry in the row is u_i or v_i. This shows that E is a GNS-cut for $\{w_1, w_2, \ldots, w_k; w'_1, w'_2, \ldots, w'_k\}$ with permutation π. This gives $\mathcal{R}_{\text{weak-edge-cut}} \supseteq \mathcal{R}_{\text{GNS-cut}}$ and completes the proof. ■

VI. k-GROUPCAST DIRECTED NETWORKS: SUM-RATE

Definition. A *k-groupcast directed network* is a $k(k-1)$-unicast directed network $\mathcal{N}$ with k distinct distinguished

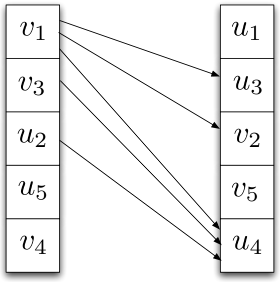


Fig. 3. One of the several 5×2 tables generated by running algorithm $\mathcal{P}$ on the $\mathcal{G}_c, \mathcal{G}_r$ shown in Fig. 2. The order of choosing sinks was v_4, u_5, $u_2v_3u_4$, u_3v_2, v_1, u_1v_5. The arrows show connectivity from the vertices on the left to the vertices on the right in $\mathcal{G}_c$. Note that arrows never go 'horizontally' or 'upward' but only go 'downward' which is the desired GNS-cut property.

nodes (group-nodes) $v_1, v_2, \ldots, v_k$ with source-destination pairs $\{s_{(i,j)}; d_{(i,j)}\}_{(i,j)\in\mathcal{I}}$ where $\mathcal{I} = \{(i,j) : 1 \leq i,j \leq k, i \neq j\}$ and $s_{(i,j)} = v_i,\ d_{(i,j)} = v_j$ for all $(i,j) \in \mathcal{I}$.

All statements in this section refer to k-groupcast directed networks.

Theorem 11. *(Naor-Zosin [17])*

$$\frac{1}{2} R_{\text{multicut}} \leq F_{\text{sum-rate}} \leq R_{\text{multicut}}. \tag{8}$$

Theorem 12.

$$C_{\text{sum-rate}} \leq 2\, R_{\text{multicut}}. \tag{9}$$

Theorems 11 and 12 together imply that routing flow is approximately capacity-achieving for sum-rate:

Corollary 13.

$$\frac{1}{2} R_{\text{multicut}} \leq F_{\text{sum-rate}} \leq C_{\text{sum-rate}} \leq 2\, R_{\text{multicut}}. \tag{10}$$

We give the proof of Theorem 12.

Proof:

Consider a k-groupcast directed network $\mathcal{N}$ with group-nodes $v_1, v_2, \ldots, v_k$. Let E be a set of edges such that $\mathcal{G} \setminus E$ has no directed paths from v_i to v_j for each $(i,j) \in \mathcal{I}$. Let $(R_{(i,j)} : (i,j) \in \mathcal{I}) \in \mathcal{C}$. Observe that E is a GNS-cut for source-destination pairs $\{s_{(i,j)}; d_{(i,j)}\}_{(i,j)\in\mathcal{I}:i>j}$. Theorem 1 gives $\sum_{(i,j)\in\mathcal{I}:i>j} R_{(i,j)} \leq \sum_{e\in E} C_e$. Similarly, we can get $\sum_{(i,j)\in\mathcal{I}:i<j} R_{(i,j)} \leq \sum_{e\in E} C_e$. Adding, we obtain $\sum_{(i,j)\in\mathcal{I}} R_{(i,j)} \leq 2\sum_{e\in E} C_e$, which completes the proof. ■

Remark 4. For the groupcast network in Fig. 1, the simple XOR coding scheme shows that the factor 2 in the statement of Theorem 12 cannot be improved upon.

VII. Discussion

It is intriguing that the kinds of symmetry that lead to flow - edge-cut closeness results also lead to the near-fundamentality of such edge-cuts. It would be interesting to find a deeper explanation of this phenomenon.

VIII. Acknowledgements

Research support for the first author from the ARO MURI grant W911NF- 08-1-0233, Tools for the Analysis and Design of Complex Multi-Scale Networks, from the NSF grant CNS- 0910702, from the NSF Science and Technology Center grant CCF-0939370, Science of Information, from Marvell Semiconductor Inc., and from the U.C. Discovery program is gratefully acknowledged.

References

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, July 2000.

[2] N. Harvey, R. Kleinberg, and A. Lehman, "On the capacity of information networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 2345–2364, June 2006.

[3] R. Dougherty, C. Freiling, and K. Zeger, "Linear network codes and systems of polynomial equations," *IEEE Transactions on Information Theory*, vol. 54, pp. 2303–2316, May 2008.

[4] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, pp. 2745–2759, August 2005.

[5] T. Chan and A. Grant, "Mission impossible: Computing the network coding capacity region," in *Proc. of IEEE ISIT*, (Toronto, Canada), July 2008.

[6] G. Kramer and S. Savari, "Edge-cut bounds on network coding rates," *Journal of Network and Systems Management*, vol. 14, pp. 49–67, March 2006.

[7] J. Chuzhoy and S. Khanna, "Polynomial flow-cut gaps and hardness of directed cut problems," in *Proc. ACM STOC*, pp. 179–188, 2007.

[8] A. Agarwal, N. Alon, and M. Charikar, "Improved approximation for directed cut problems," in *Proc. of ACM STOC*, pp. 671–680, 2007.

[9] F. Leighton and S. Rao, "An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms," in *Proc. of 28th Annual Symposium on Foundations of Computer Science*, (Los Alamitos, California), 1988.

[10] N. Linial, E. London, and Y. Rabinovich, "The geometry of graphs and some of its algorithmic applications," *Combinatorica*, vol. 15, no. 2, pp. 215–245, 1995.

[11] S. Arora, J. Lee, and A. Naor, "Euclidean distortion and the sparsest cut," *J. of the AMS, to appear, Preliminary version in Proc. of ACM STOC*, 2005.

[12] Z. Li and B. Li, "Network coding: The case of multiple unicast sessions," in *Proc. of the 42nd Allerton Annual Conference on Communication, Control and Computing*, September 2004.

[13] N. Harvey, R. Kleinberg, and A. Lehman, "Comparing network coding with the multicommodity flow for the k-pairs communication problem," in *MIT LCS, Tech.Rep. 964*, September 2004.

[14] P. Klein, S. Plotkin, S. Rao, and E. Tardos, "Bounds on the max-flow min-cut ratio for directed multicommodity flows," *J. Algorithms*, no. 22, pp. 241–269, 1997.

[15] S. Kamath, D. Tse, and V. Anantharam, "Generalized network sharing outer bound and the two-unicast problem," in *Proc. of International Symposium on Network Coding*, (Beijing, China), July 2011.

[16] I. Wang, S. Kamath, and D. Tse, "Two unicast information flows over linear deterministic networks," in *Proc. of IEEE ISIT*, (St. Petersburg, Russia), August 2011.

[17] J. Naor and L. Zosin, "A 2-approximation algorithm for the directed multiway cut problem," *SIAM Journal on Computing*, vol. 31, no. 2, pp. 477–482, 2001.

[18] C. Chekuri, S. Kannan, A. Raja, and P. Viswanath, "Multicommodity flows in polymatroidal capacity networks." http://www.ifp.illinois.edu/~pramodv/pubs/techreport.pdf, 2011.

[19] S. Kamath, S. Kannan, and P. Viswanath, "Wireless networks with symmetric demands," in *Proc. of IEEE ISIT*, (Cambridge, MA, USA), July 2012.

[20] M. Saks, A. Samorodnitsky, and L. Zosin, "A lower bound on the integrality gap for minimum multicut in directed networks," *Combinatorica*, vol. 24, no. 3, pp. 525–530, 2004.