Cut-set Theorems for Multi-state Networks

Mohammad Ali Khojastepour, Ashutosh Sabharwal, and Behnaam Aazhang

Abstract

We derive new cut-set bounds on the achievable rates in a general multi-terminal network with finite number of states. Multiple states are common in communication networks in the form of multiple channel and nodes' states. Our results are broadly applicable and provide much tighter upper bounds than the known single state min-cut max-flow theorem, and hence form an important new tool to bound the performance of multi-node networks.

Two examples are presented to illustrate the tightness and the utility of the new bounds. In each of the example applications, the known single-state max-flow min-cut theorem provides a bound strictly looser than the new cut-set bounds. The first illustrative example is single-user compound channels, where both the transmitter and receiver have channel state information. The example of compound channels represents the smallest possible network with only two nodes, but has multiple states due to channel variations. The upper bound derived using the proposed bounds turns out to be the capacity of the compound channel, which implies that the bound is tight in this case. The second example is from a contemporary network problem. Here, we demonstrate the application of new bounds to characterize the limits on rate of information transfer in 'cheap' relay networks, where the relay nodes can either transmit or receive, but not both simultaneously. In this case, each constituent channel has a single state but relay nodes can be in one of the two states, transmit or receive mode, giving rise to multiple network states. Here, again, the upper bound coincide with the capacity of the channel if the relay channel is degraded.

Index Terms

Network information theory, capacity, achievable rates, coding theorem, converses.

Mohammad Ali Khojastepour, Ashutosh Sabharwal, and Behnaam Aazhang are in the Department of Electrical and Computer Engineering, Rice University, Houston, TX.

I. INTRODUCTION

Network information theory studies the problems related to reliable communication between subsets of nodes in a given network of senders and receivers. Finding limits on the reliable communication in a multi-terminal network often proves to to be very challenging, primarily due to simultaneous existence of interference, cooperation and feedback between the nodes. The challenge facing network information theory is evident from the fact that the exact capacity is known for only a handful of problems. Hence, we often rely on lower and upper bounds to understand the limits of communication efficiency. One of the most powerful and versatile bound is known as min-cut max-flow theorem [1]. Interestingly, in most cases where the capacity of a network is known, the min-cut max-flow theorem [1] coincides with the achievable rates, making it both a useful bound and a method to prove converses to many capacity theorems.

Due to the difficulty of the problems in network information theory, it is commonly believed that the problem of finding the capacity of an arbitrary network will remain unsolved. In light of the above challenges, the general upper bound in [1] has witnessed more increased attention to understand network capacity limits. For example, cut-set bounds for single-state networks were used in [2] to study the asymptotic behavior in large scale networks when there is only one active flow. The cut-set bounds were also used in [3] to study a variation of multiple relay problem. However, when the min-cut max-flow theorem is applied directly to networks with multiple states the resulting bound is loose to the extent that it is not very useful for either understanding the limits of the network or engineering the nodes.

In this paper, we derive cut-set theorems for multi-terminal networks with possibly multiple states. The multi-state min-cut max-flow theorem reduces to the well-known min-cut max-flow theorem [1] when applied to networks with a single state. In this paper, we show that if the network has multiple states, the new theorem provides a tighter upper bound than the bound in [1].

A network can have multiple states for many possible reasons. Commonly studied possibilities include channels and sources with multiple states, of which we discuss the example of compound channel capacity in a single-user system. Our result also applies to the cases where the transmission and reception of nodes is restricted because of the hardware limitations of each node. As an example, we establish the capacity of a 'cheap' relay channel [4–7], where the relay node can either transmit or receive but not both simultaneously.

The results in this paper are derived for the case of discrete memoryless networks with finite number of states. Though the assumption of memoryless is fundamental to derivation of new bounds, the assumptions of discrete alphabet and finite number of states can be relaxed in many cases of interest. For example, the bounds of this paper was extended to Gaussian 'cheap' relay channel by the authors in [6], which is a continuous alphabet channel. An example extension to infinite number of states are the bounds on capacity of fading channels derived in $[8, 9]$.

The rest of the paper is organized as follows: We formally define the problem and notations in Section II. The main results which is the cut-set bounds for the multi-state network is derived in Section III. In Section IV, we show some of the applications of the derived bounds. Finally, we conclude in Section V.

II. Problem Formulation

We consider a general multi-terminal network of senders and receivers, shown in Figure 1, where the network has finite number of states. The network can be considered as a directed graph where each node of the network represents a potential sender or receiver, and each link represents existence of a one-way channel between two nodes. We define a *network channel* to denote both the full network and a subset of nodes in the network, and distinguish it from a channel which exists between any two nodes. Alternately, the network channel can be understood as a collection of individual channels between the nodes in a network.[1] In accordance with the term *channel use* in a single channel, we define *network use* which corresponds to the simultaneous one time use of the network channel in a given state.

Assume that there are N nodes and for every node i, the input random variable $X^{(i)}$ is the transmitted signal, and the output random variable $Y^{(i)}$ is the received signal. We assume that all intended messages $W^{(ij)}$ from node i to the node j, which are going to be transmitted in n network uses, are independent and uniformly distributed over their respective ranges $\{1, 2, 3, ..., 2^{nR^{(ij)}}\}$, where $R^{(ij)}$ represents the rate in which node i transmits information to node j.

Considering a network with finite number of states and following the notation of [11], the channel is represented by the collection of channel transition functions or channel probability

[1] The term *network channel* was first introduced in [10], where the reader can find a more complete discussion.

functions (c.p.f.) defined[2] as: $p(y^{(1)}, y^{(2)}, \ldots, y^{(N)}|\ x^{(1)}, x^{(2)}, \ldots, x^{(N)}|m)$ where m is the state of the network which takes its values from a set of possible states, $\mathcal{M}$, with finite cardinality $M = |\mathcal{M}|$.

The input symbol $X_k^{(i)}$ which is the transmitted signal from node i in time k (or k^{th} network use) depends on $W^{(ij)}\ \forall j \in \{1, 2, \ldots, N\}$, and also the past values of the received signal $Y^{(i)}$, i.e., $Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_{k-1}^{(i)}$. Also, node j decodes the intended message $W^{(ij)}$ for this node from node i, $\forall i \in \{1, 2, \ldots, N\}$ after the transmission of the whole block of length n. Therefore, decoding of this message is based on all the received signals $Y_1^{(j)}, Y_2^{(j)}, \ldots, Y_n^{(j)}$ and the transmitted massage from this node, i.e., $W^{(jl)}\ \forall l \in \{1, 2, \ldots, N\}$. Thus, the encoding and decoding functions of block length n code for node i have the following structure:

Encoder: $X_k^{(i)}(W^{(i1)}, W^{(i2)}, \ldots, W^{(iN)}, Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_{k-1}^{(i)})$ for any node $i \in \{1, 2, \ldots, n\}$ and any network use $k \in \{1, 2, \ldots, N\}$.

Decoder: $\widehat{W}^{(ij)}\left(Y_1^{(j)}, Y_2^{(j)}, \ldots, Y_n^{(j)}, W^{(j1)}, W^{(j2)}\ , \ \ldots, W^{(jN)}\right)$ for all values of $i, j \in \{1, 2, \ldots, N\}$ which is estimate of the receiver of node i at node j based on the received signals at node j for the whole block of n transmissions and its own transmitted information for the other nodes.

The probability of error for each decoder is defined as: $P_e^{(n)(ij)} = \text{Prob}(\widehat{W}^{(ij)} \neq W^{(ij)})$ which is defined based on the assumption that the messages are independent and uniformly distributed over their respective ranges.

A set of rates $\{R^{(ij)}\}$ is said to be achievable if there exist an encoding and decoding function with block length n such that $P_e^{(n)(ij)} \longrightarrow 0$ when $n \longrightarrow \infty$ for all $j, i \in \{1, 2, \ldots, N\}$.

We denote the state of the channel in the k^{th} network use as m_k. For any state m define $n_m(k)$ as the number of times that the network is used in state m in the first k network uses. Let

$$t_m = \lim_{k \longrightarrow \infty} \frac{n_m(k)}{k} \tag{1}$$

define the portion of the time that network has been used in state m as the total number of network uses goes to infinity.

[2]The notation $p(\cdot|\cdot|\cdot)$ is unconventional, but was used in [11] to represent channel probability functions as a function of the channel state

III. Multi-State Cut-Set Bounds

In this section, we first derive a bound on the sum rate of the information transfer across one cut-set in the multi-state network (Section III-A). This forms a basis for our main result which is the max-flow min-cut theorem for the multi-state network (Section III-B). The results of this section are derived assuming that the inputs and outputs are discrete and the network channel is memoryless. Though the memoryless property of the channel is vital in our derivation, it is relatively easy to extend the same results for continuous input alphabets [6]. It should be noted that in the continuous alphabet case, typically additional constraints on the input signals (such as limited average power, or peak power) are needed in order to have finite capacity, which requires suitable modifications to subsequent analysis.

A. Single Cut Bound

We partition the given network into two subsets, set $S \subset \{1, 2, \ldots, N\}$ and its complement, S^c. We will derive two bounds on the rate of information flow from senders in Set S to receivers in Set S^c. When the sequence of the network states $(m_k)_{k=1}^{\infty} = (m_1, m_2, m_3, \ldots)$ are known a priori, Lemma 1 provides the first upper bound on the achievable rates of transmission across a give cut-set. And in the second bound (Theorem 1), we consider the case where the sequence of the network states can be chosen arbitrarily.

Lemma 1 (Single Cut Bound): Consider a general network with finite number of states, M, for which the sequence $(m_k)_{k=1}^{\infty}$ of the network states is given and known a priori to all nodes. If the information rates $\{R^{(ij)}\}$ are achievable then there exist some joint probability distributions $p(x^{(1)}, x^{(2)}, \ldots, x^{(N)}|m)$, $m = 1, 2, \ldots, M$, such that for any $S \subset \{1, 2, \ldots, N\}$,

$$\sum_{i \in S, j \in S^{(c)}} R^{(ij)} \leq \sum_{m=1}^{M} t_m I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \tag{2}$$

where S^c denotes the complement of the set S, $I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c})$ is the conditional mutual information between $X_{(m)}^{S}$ and $Y_{(m)}^{S^c}$ given $X_{(m)}^{S^c}$, and $\{t_m\}_{m=1}^{M}$ represents the asymptotic proportions of the states.

Proof: See Appendix I. ■

It should be noted that this bound is not a function of the choice of the sequence of the network states, $(m_k)_{k=1}^{\infty}$, directly, rather it depends only on the asymptotic values $t_1, t_2, ...t_M$. In fact this bound holds for the best choice of the deterministic function[3] m_k, as well as the case where it changes randomly from one network use to the other while the asymptotic proportions of the sequence of the states, i.e., $t_i, i \in \{1, 2, \ldots, M\}$ remain fixed. Thus, no information can be embedded in the order of states m_k if their asymptotic proportions are held constant.

If the sequence of the network states can be chosen arbitrarily, the bound across each cut-set would be the supremum of the achievable bounds in Lemma 1 for all choices of the function m_k. On the other hand, since the bound is not a function of the individual sequence of the network states, the supremum would be taken over the possible values of $t_i, i \in \{1, 2, \ldots, M\}$ and specifically we have the following theorem.

Theorem 1: Consider a general network with finite number of states, M. Maximum achievable information rates $\{R^{(ij)}\}$ across the cut-set $S \subset \{1, 2, \ldots, N\}$ for any choice of network state sequence $(m_k)_{k=1}^{\infty}$ is bounded by:

$$\begin{aligned} \sum_{i \in S, j \in S^{(c)}} R^{(ij)} &\leq \sup_{t_m} \sum_{m=1}^{M} t_m I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \\ &= \sup_{m} I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \end{aligned} \tag{3}$$

for some joint probability distributions $p(x^{(1)}, x^{(2)}, \ldots, x^{(N)} | m)$, $m = 1, 2, \ldots, M$, where $\sum_{i=1}^{M} t_m = 1$.

Thus, in order to maximize the sum rate of the information transfer across a cut-set, the above result suggests using the network just in one fixed state m which allows for the maximum mutual information $I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c})$. Though Theorem 1 is a bound on actual capacity, intuitively it appears that the actual capacity will have similar behavior of only using the most beneficial state of the network. On the other hand, in the case that the sequence of state cannot be chosen, average mutual information is the bound which is given by Lemma 1. In Section IV, we show the connection of the above bound to the results on compound channel capacity.

[3]Section IV shows an interesting case of a deterministic function where this new cut-set bound is tight.

B. Max-Flow Min-Cut Bound

In many cases, the aim is to maximize the information rate between two disjoint subsets of nodes in the network regardless of any possible simultaneous communication rates between other nodes in the network. For example, this situation happens in the relay problem when a source node (or set of source nodes) transmits information to a destination node (or set of destination nodes) by the help of some intermediate relay node (or relay nodes). In this case, the aim is to maximize the rate of information transfer from source to destination. Therefore, instead of a bound on a single cut-set in the network, we would like to find an upper bound between the rate of information transfer between two disjoint subsets of nodes $S_1, S_2 \subset \{1, 2, \ldots, N\}$ of the network. For any cut-set that partitions the set of nodes in the network to $\{S, S^{(c)}\}$ and has either one of the sets S_1 and S_2 in just one side of the cut (*e.g.*, in Figure 1 consider sets S_1 and S_2 for either cuts C_1 or C_2), we can use the single cut bounds of Section III-A. Therefore, the rate of information transfer between S_1 and S_2 is bounded by the minimum of all such single cut bounds.

A simple observation suggests that the sum of information rates from the set S_1 to the set S_2 is bounded by the sum of the rates across all the cut-sets S such that $S \subset \{1, 2, \ldots, N\}$, $S \cap S_1 = S_1$, $S \cap S_2 = \phi$ which is given by Theorem 1. Thus, we have:

$$\sum_{i \in S_1, j \in S_2} R^{(ij)} \leq \min_S \sup_m I(X^S_{(m)}; Y^{S^c}_{(m)} | X^{S^c}_{(m)}) \tag{4}$$

for some joint probability distribution $p(x^{(1)}, x^{(2)}, \ldots,\ x^{(N)} | m)$.

The bound of (4) is a naive bound, because it does not capture the fact that the sequence of network states which gives the maximum achievable sum rate across a given cut-set does not necessarily leads to the maximum achievable rate across other cut-sets in the network. Therefore, since the maximum achievable rates across different cut-sets depend on the sequence of the network states, we use Lemma 1 directly to find better upper bound for the achievable rate of the information flow between the two disjoint subsets of the nodes $\{S_1, S_2\}$.

Lemma 2: Consider a general network with finite number of states, M, for which the sequence $(m_k)_{k=1}^{\infty}$ of the states of the network is given and known to all nodes. If the information rates $\{R^{(ij)}\}$ are achievable then the sum rate of information transfer from a node set S_1 to a disjoint node set S_2 where $S_1, S_2 \subset \{1, 2, \ldots, N\}$, is bounded by:

$$\sum_{i \in S_1, j \in S_2} R^{(ij)} \leq \min_{S} \sum_{m=1}^{M} t_m I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c}), \tag{5}$$

for some joint probability distributions $p(x^{(1)}, x^{(2)}, \ldots, x^{(N)}|m)$, $m = 1, 2, \ldots, M$, where the minimization is taken over all set $S \subset \{1, 2, \ldots, N\}$ subject to $S \cap S_1 = S_1$, $S \cap S_2 = \phi$.

Proof: Direct application of the Lemma 1 over multiple cut-sets, and considering the fact that the sequence of network state is given, results in (5). ■

Again, considering all possible sequences of network states, the bound of the achievable rate of information transfer from a node set S_1 to a disjoint node set S_2, where $S_1, S_2 \subset \{1, 2, \ldots, N\}$, is given by the supremum of the achievable bounds in Lemma 2 for all choices of network state sequence function m_k. It again translates to taking supremum over all possible asymptotic values, $t_i, i \in \{1, 2, \ldots, M\}$.

Theorem 2: Consider a general network with finite number of states, M. If the information rates $\{R^{(ij)}\}$ are achievable then the sum rate of information transfer from a node set S_1 to a disjoint node set S_2, where $S_1, S_2 \subset \{1, 2, \ldots, N\}$, and for any choice of network state sequence $(m_k)_{k=1}^{\infty}$ is bounded by:

$$\sum_{i \in S_1, j \in S_2} R^{(ij)} \leq \sup_{t_m} \min_{S} \sum_{m=1}^{M} t_m I(X_{(m)}^{S}; Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \tag{6}$$

for some joint probability distributions $p(x^{(1)}, x^{(2)}, \ldots, x^{(N)}|m)$, $m = 1, 2, \ldots, M$, when the minimization is taken over all set $S \subset \{1, 2, \ldots, N\}$ subject to $S \cap S_1 = S_1$, $S \cap S_2 = \phi$ and the supremum is over all the non-negative t_m subject to $\sum_{i=1}^{M} t_m = 1$.

In the next section we will consider some specific examples of the networks and application of this bound on the achievable rates of communication in such networks.

IV. Applications of the New Bounds

When there is only one state of operation in the network, i.e., $(M = 1)$, the bound of Lemma 1 reduces to the known cut-set theorem in network information theory [Theorem 14.10.1 in [1]], and Theorem 2 can be interpreted as max-flow min-cut theorem for a single state network with $(M = 1, S_1 = S, S_2 = S^{(c)})$. Although these cut-set bounds are not tight in general, they are tight enough to find the capacity region or capacity in some cases such as multiple access channel, degraded relay channel, or general relay channel with feedback [12].

In the sequel, we will discuss two applications of the new multi-state cut-set bounds. First, we will consider the application of Lemma 1 to derive an upper bound on the capacity of compound channels (Section IV-A). For this example, the network is trivial and has only one channel, i.e., between the sender and receiver nodes. Multiple states in the network are due to changes in channel probability function from one channel use to another. In this case, the upper bound from Lemma 1 turns out to be the capacity of compound channels.

The second application is an upper bound for the capacity of 'cheap' relay channels (Section IV-B), for which previously known cut-set bound does not provide a tight bound. However, the derived bounds in this paper not only provide a better bound but tight for the case of degraded 'cheap' relay channel.

A. Compound Channel

Channels with finite number of states were considered first in the context of compound channels [11], and the capacity expression for discrete memoryless channels with different cases of channel state information at the receiver and the transmitter have been derived. Extension of the results in [11] for the fading channel which is a multi-state channel with infinite number of states can also be found in [8, 9]. The derived cut-set bounds for the multi-state networks unifies some of the converses for the known multi-state channels. For example, Lemma 1 provides a tight bound for the capacity of the compound channel with the known channel state information both at the transmitter and receiver (Theorem 4.6.1 in [11]).

Compound channel is defined as a channel for which the channel probability function which governs the transmission of signal changes with time, but independently of all previously transmitted and received signals, with a known probability distribution $P_{state}(m)$. Lemma 1 provides a very simple upper bound for the capacity of the compound channel, $C_{compound}$, where noncausal channel state information is available to both transmitter and receiver. We have

$$C_{compound} \leq \sum_{m=1}^{M} t_m I(X_{(m)}; Y_{(m)}) = \sum_{m=1}^{M} t_m C(m) \tag{7}$$

where $C(m)$ represents the channel capacity of a uni-state channel in state m. In fact, it has been shown that this upper bound is the exact capacity of the compound channel (Theorem 4.6.1 in [11]), thus the bound of Lemma 1 is tight in this case.

It can also be shown directly that the average mutual information bound of Lemma 1 is an upper bound to the capacity of the channel with noncausal side information at the transmitter introduced in [13], although in this case the bound is loose in general. We have

$$\max_{p(u,x|v)} \{I(U;Y) - I(U;V)\} \leq \sum_{m=1}^{M} t_m I(X_{(m)}; Y_{(m)}), \tag{8}$$

where V represents the channel state which is available to the transmitter for the whole block of transmission a priori, and U is an auxiliary random variable with a finite alphabet size.

Two notes are in order here. First, the bounds in this paper are easily extendable for channels with continuous input and output alphabets. For example, Lemma 1 can also serve to find a bound on the result known as "writing on dirty paper" which can be considered as an extended version of the channel with noncausal side information at the transmitter [13] with continuous alphabet.

Second, although the bounds in this paper are obtained with the assumption of finite number of states in the network, they can also be applied to some channels with infinite number of states. For example, Lemma 1 provides tight upper bound in (i) the fading channel with side information and average power constraint [8], and (ii) the fading channel with side information and both peak and average power constraint [9]. The fading channel with side information [8, 9] can be considered as an extended version of the compound channel [11] with continuous alphabet and infinite number of states.

B. 'Cheap' Relay Channel

A *cheap* node is defined as a node which is either a sender or receiver for any channel use. Each state of operation in a network with cheap nodes corresponds to the valid partitioning of the nodes into two disjoint subsets of sender and receiver nodes. It can be deduced that there are finite number of states for a network with finite number of nodes.

The 'cheap' relay channel, Figure 3, is a variation of the relay channel where the relay node is either transmitter or receiver at any given channel use (or equivalently any given time). The 'cheap' relay channel is introduced in [4] and an achievable rate for the defined channel is derived.

Let $p(y, y_1|x_1, x_2)$ denote the channel probability function of the original relay channel. Since the relay node, R, is either transmitter or receiver there are two possible state of operation in the network. In the state m_1 ("broadcast state" in Figure 3), the relay node R acts as a receiver and thus channel probability function is given by $p(y_1|x_1|m_1) = p(y, y_1|x_1, x_2 = 0)$. While in the state m_2 ("Multiple access state" in Figure 3), the relay node functions as a transmitter and the channel probability function is given by $p(y|x_1, x_2|m_2) = \sum_{y_1} p(y, y_1|x_1, x_2)$. We use $x_2 = 0$ to denote the case that there is no input from relay node, R, to the channel in state m_1. It should be pointed out that we use the terms "broadcast state" and "multiple access state" because of the resemblance of the network in these states with broadcast and multiple access channel.

However, to achieve the full capacity of the channel coding scheme cannot be divided into two separate coding scheme one for each individual state. For example, consider the case that the received signal at the destination is degraded version of the received signal at the relay in broadcast state. In this situation, the destination cannot decode the whole message which is sent in this state whereas it is decodable at the relay node. It means that there would be an ambiguity about the whole or at least part of the transmitted message from the source which remains to be resolved by next transmission in multiple access state. In this case, resolving the ambiguity is easier than retransmission of the whole message in multiple access state. In other words, although the received signal at the destination node in broadcast state is not enough to decode whole transmitted message, it would provide a partial information that would be resolved later in multiple access state.

From Theorem 2, an upper bound for the information transfer rate R from source node S to the destination node D would be

$$\begin{aligned} R \leq \sup_{t,\, 0 \leq t \leq 1} \min\{ & t\, I(X_1; Y, Y_1|m_1) + (1-t)\, I(X_1; Y| \\ & X_2, m_2), t\, I(X_1; Y|m_1) + (1-t)\, I(X_1, X_2; Y|m_2)\}. \end{aligned} \tag{9}$$

where there is minimum of two terms is involved: the first term corresponds to the bound on broadcast cut-set C_1 and the second term corresponds to the multiple access cut-set C_2

in Figure 2. Either of the above cut-set bounds over C_1 and C_2 have two conditional mutual information terms. The first conditional mutual information term corresponds to the broadcast state, m_1, and the second conditional mutual information corresponds to the multiple access state.

On the other hand, it has been shown [4] that for every t, $0 \leq t \leq 1$, the rate R^* is achievable where R^* is given by

$$\begin{aligned} R^* \triangleq \min\{& t\, I(X_1;Y_1|m_1) + (1-t)\, I(X_1;Y|X_2,m_2), \\ & t\, I(X_1;Y|m_1) + (1-t)\, I(X_1,X_2;Y|m_2)\}. \end{aligned} \tag{10}$$

Thus, if in the state m_1 the received signal y at the destination node D is degraded form of the received signal y_1 at the relay node then the bound of (9) would coincide with this achievable rate for some value of t. Hence, the bound derived in Theorem 2 provides the converse for the capacity theorem of degraded cheap relay channel, and the capacity is given by $C = R^*$ defined in (10).

It should be pointed out that the achievable rate of information transmission in the cheap relay channel is lower than that of the relay channel without the extra condition on relay node (Figure 4). Thus, the known cut-set theorem of network information theory is unable to provide tight enough bounds. On the other hand, the derived bounds in this paper prove to be effective in this case and they capture the effect of the practical constraint imposed on the network nodes to the extent that we can actually drive the capacity of degraded cheap relay channel.

Although having more than one state of operation in the network seems to increase degrees of freedom and thus the achievable rates in the network, it is not always true. For example, in the problem of communication in a network with cheap nodes [4], having more than one state is a result of imposing the condition that each node cannot transmit and receive at the same time. It is somehow trivial that posing this condition will not increase any sets of achievable rates in the network in comparison to the same network without this constraint (which has just one state of operation but all of the links or channels between the nodes can be used at the same time).

V. Concluding Remarks

Considering the fact that cut-set bounds are tight in many important cases (multiple access, relay channels, etc.), they have received relatively little attention in the literature. The reason may partly lie in the fact that point-to-point (two node network) analysis received far greater attention than other problems, especially those in network information theory. We envision that the situation will gradually change as the focus of the community shifts to analyzing capacity of ad hoc, sensor and distributed wireless networks.

We acknowledge that the new bounds in this paper are only a step towards capturing the complexities of actual networks. To enable significant advances, we feel that new methods which address the following two challenges are critical. First, the cut-set bounds inherently assume a collaborative reception by the nodes in the receive subset (*e.g.*, set S_2 in Figure 1). This is precisely the reason why min-cut max-flow cut-set theorem does not yield the capacity of broadcast channels, where the receivers decodes their signals with no collaborative communication with other receivers.

Second, a more serious deficiency of known cut-set bounds is that they provide no useful bound if there are loops in the network; a classic example of this situation is the two-way channel. Feedback via shared bandwidth is an important aspect of all networks, and information theoretic methods have had limited success in incorporating in its analysis [14].

Appendix I

Proof of Lemma 1: Let $T = \{(i,j) : i \in S, j \in S^c\}$ be the set of all links that cross from node set S to the node set S^c and T^c be all the other links in the network. Since $W^{(ij)}$ are independent and uniformly distributed over their range and based on the definitions and conditions of Section II, we have

$$\begin{aligned}
&\sum_{i\in S, j\in S^{(c)}} R^{(ij)} \\
&= \sum_{i\in S, j\in S^{(c)}} H\left(W^{(ij)}\right) && (11)\\
&= H\left(W^{T}\right) && (12)\\
&= H\left(W^{T} \left| W^{T^c}\right.\right) && (13)
\end{aligned}$$

$$
\begin{aligned}
&= I\left(W^T; Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_N^{(S^c)} \middle| W^{(T^c)}\right) \\
&\quad + H\left(W^T \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_N^{(S^c)}, W^{(T^c)}\right) && (14) \\
&\leq I\left(W^T; Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_N^{(S^c)} \middle| W^{(T^c)}\right) + n\epsilon_n && (15) \\
&= \sum_{k=1}^{N} I\left(W^T; Y_k^{(S^c)} \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_{k-1}^{(S^c)}, W^{(T^c)}\right) \\
&\quad + n\epsilon_n && (16) \\
&= \sum_{k=1}^{N} \left\{ H\left(Y_k^{(S^c)} \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_{k-1}^{(S^c)}, W^{(T^c)}\right) \right. \\
&\quad \left. - H\left(Y_k^{(S^c)} \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_{k-1}^{(S^c)}, W^{(T^c)}, W^T\right) \right\} \\
&\quad + n\epsilon_n && (17) \\
&\leq \sum_{k=1}^{N} \left\{ H\left(Y_k^{(S^c)} \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_{k-1}^{(S^c)}, W^{(T^c)}, \right. \right. \\
&\quad \left. X_k^{(S^c)}\right) - H\left(Y_k^{(S^c)} \middle| Y_1^{(S^c)}, Y_2^{(S^c)}, \ldots, Y_{k-1}^{(S^c)}, \right. \\
&\quad \left. \left. W^{(T^c)}, W^T, X_k^{(S^c)}, X_k^{(S)}\right) \right\} + n\epsilon_n && (18) \\
&\leq \sum_{k=1}^{N} \left\{ H\left(Y_k^{(S^c)} \middle| X_k^{(S^c)}\right) - H\left(Y_k^{(S^c)} \middle| X_k^{(S^c)}, X_k^{(S)}\right) \right\} \\
&\quad + n\epsilon_n && (19) \\
&= \sum_{k=1}^{N} I\left(X_k^{(S)}; Y_k^{(S^c)} \middle| X_k^{(S^c)}\right) + n\epsilon_n && (20) \\
&= \sum_{k=1}^{N} I\left(X_{Q_{m(k)}}^{(S)}; Y_{Q_{m(k)}}^{(S^c)} \middle| X_{Q_{m(k)}}^{(S^c)}, Q_{m(k)} = k\right) \\
&\quad + n\epsilon_n && (21) \\
&= \sum_{m=1}^{M} n_m(n) I\left(X_{Q_m}^{(S)}; Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}, Q_m\right) + n\epsilon_n && (22) \\
&= \sum_{m=1}^{M} n_m(n) \left\{ H\left(Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}, Q_m\right) \right. \\
&\left. - H\left(Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}, X_{Q_m}^{(S)}, Q_m\right) \right\} + n\epsilon_n && (23) \\
&\leq \sum_{m=1}^{M} n_m(n) \left\{ H\left(Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}\right) \right.
\end{aligned}
$$

$$-H\left(Y_{Q_m}^{(S^c)}\left|X_{Q_m}^{(S^c)},X_{Q_m}^{(S)}\right.\right)\Big\}+n\epsilon_n \tag{24}$$

$$=\sum_{m=1}^{M}n_m(n)I\left(X_{Q_m}^{(S)};Y_{Q_m}^{(S^c)}\left|X_{Q_m}^{(S^c)}\right.\right)+n\epsilon_n \tag{25}$$

where

(11) follows from the assumption that $W^{(ij)}$'s are distributed uniformly over their respective ranges $\left\{1,2,3,...,2^{nR^{(ij)}}\right\}$,

(12) follows from the assumption that $W^{(ij)}$'s are independent and also definition $W^{(T)}=\left\{W^{(ij)}:i\in S,j\in S^c\right\}$,

(13) follows from the independence of $W^{(T)}$, $W^{(T^c)}$,

(14) follows from the definition of the mutual information,

(15) follows from the Fano's inequality, because message $W^{(T)}$ can be decoded from $Y^{(S)}$ and $W^{(T^c)}$. (Since we have assumed that the set of rates $\left\{R^{(ij)}\right\}$ are achievable; and also note that $\epsilon_n\longrightarrow 0$ as $n\longrightarrow\infty$),

(16) follows from the chain rule,

(17) follows from the definition of the mutual information,

(18) follows from the fact that first term has been changed due to the definition of $X_k^{(S^c)}$ which is function of the past received symbols $Y(S^c)$ and the message $W^{(T^c)}$. Also the second term has not been increased since conditioning can only reduce the entropy,

(19) follows from the fact that $Y_k^{(S^c)}$ only depends on the current input symbols $X_k^{(S^c)}$, $X_k^{(S)}$. (Note that although $Y_k^{(S^c)}$ depends on the state m_k but it is deterministic and predefined before transmission for all the nodes),

(20) follows from the definition of the mutual information,

(21) follows from introducing time sharing random variables $Q_1,Q_2,\ldots,Q_M$ for each state of the network, where each state m, Q_m is uniformly distributed over all values of the time index k, $k\in\{1,2,3,...,N\}$, for which $m_k=m$,

(22) follows from rearranging the summation and using definition of the average mutual information,

(23) follows from the definition of the mutual information,

(24) follows from the fact that conditioning cannot increase the entropy for the first term and

the fact that second term has not been changed since $Y_{Q_m}^{(S^c)}$ only depends on the current input symbols $X_{Q_m}^{(S^c)}$, $X_{Q_m}^{(S)}$ and conditionally is independent of Q_m,
(25) follows from the definition of the mutual information.

Thus, by dividing both sides of inequality (25) by n and finding the limit as $n \longrightarrow \infty$ we have

$$\sum_{i \in S, j \in S^{(c)}} R^{(ij)} \leq \sum_{m=1}^{M} \frac{n_m(n)}{n} I\left(X_{Q_m}^{(S)}; Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}\right) + \epsilon_n. \tag{26}$$

As $n \longrightarrow \infty$ we have:

$$\sum_{i \in S, j \in S^{(c)}} R^{(ij)} \leq \sum_{m=1}^{M} t_m I\left(X_{Q_m}^{(S)}; Y_{Q_m}^{(S^c)} \middle| X_{Q_m}^{(S^c)}\right) \tag{27}$$

and it completes the proof. ■

REFERENCES

[1] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[2] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *IEEE Infocom Conference on Computer Communications*, New York City, 2002.

[3] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proceedings of the IEEE International Symposium on Information Theory*, 2000.

[4] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of 'cheap' relay networks," in *Proceedings of 37th Annual Conference on Information Sciences and Systems*, Baltimore, Maryland, 2003, pp. 12–14.

[5] ——, "Achievable rates for arbitrary network topologies with 'cheap' nodes," in *Proceedings of the 2003 Canadian Workshop on Information Theory (CWIT-2003)*, Waterloo, Ontario, May 2003, pp. 18–21.

[6] ——, "On capacity of Gaussian 'cheap' relay channel," in *Proceedings of IEEE 2003 Global Communications Conference (Globecom-2003)*, 2003.

[7] ——, "Wireless networks with 'cheap' nodes: Achievable rates and capacity theorems," *In prepration for the IEEE Trans. on Information Theory*.

[8] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[9] M. A. Khojastepour and B. Aazhang, "The capacity of average and peak power constrained fading channel with channel side information," in *Submitted to IEEE-Wireless Communications and Networking Conference*, 2004.

[10] A. Khoshnevis and A. Sabharwal, "Network channel estimation in cooperative wireless networks," in *Proceedings Canadian Workshop on Information Theory*, 2003.

[11] J. Wolfowitz, *Coding Theorems of Information Theory*. Springer-Verlog, Berlin Heidelberg Newyork, 1978.

[12] T. M. Cover and A. E. Gamal, "Capacity thorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584,, Sept. 1979.

[13] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Problem of Control and Information Theory*, vol. 9, no. 1, 1980.

[14] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummated union," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2416–2434, October 1998.

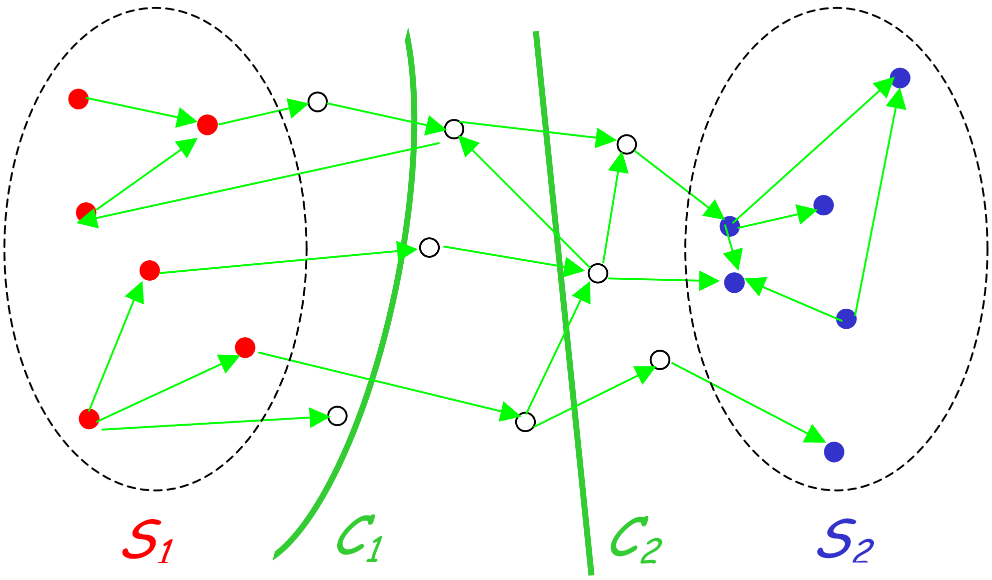

Fig. 1. A general multi-terminal network. Two arbitrary cut-sets C_1 and C_2 are shown, each of which have the set of source nodes, S_1, on one side and the set of destination nodes, S_2, on the other side.

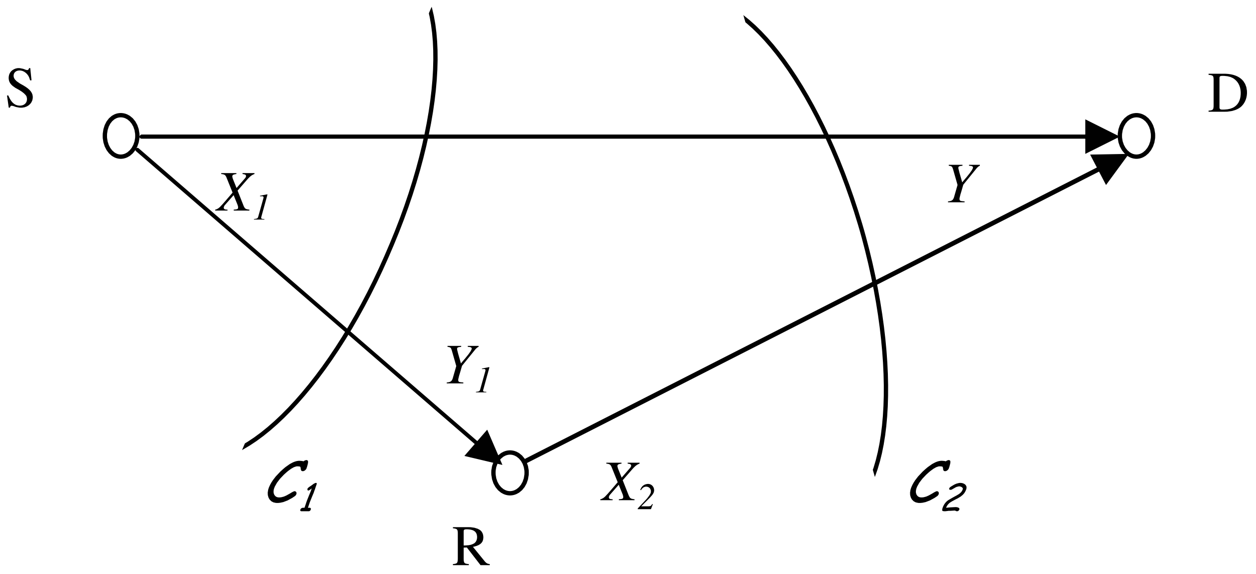

Fig. 2. 'Cheap' Relay Channel: Representation of the cut-sets. Notes: 1-The relay can either transmit or receive at any given time. 2-Both of the cut-sets exist in either of Broadcast or Multiple access states.

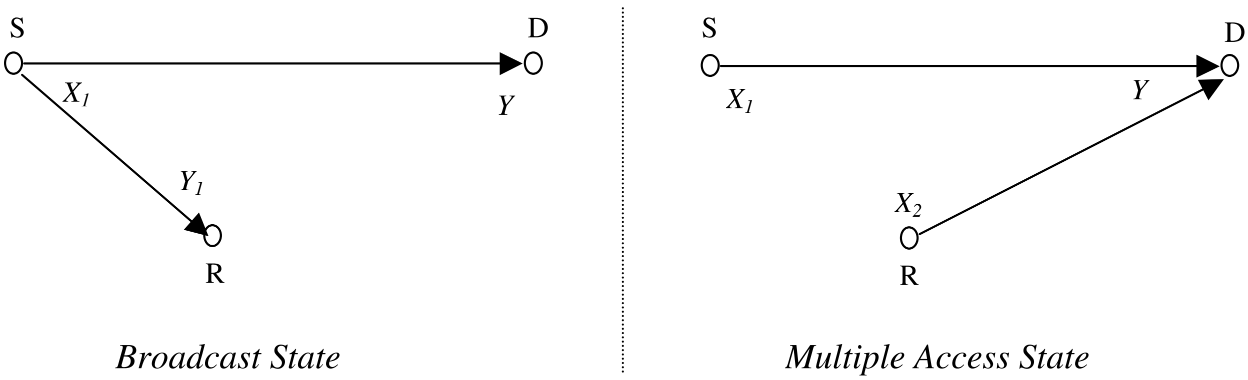

Fig. 3. 'Cheap' relay channel: The representation of two states of operation as "broadcast state" (m_1) and "multiple access state" (m_2)

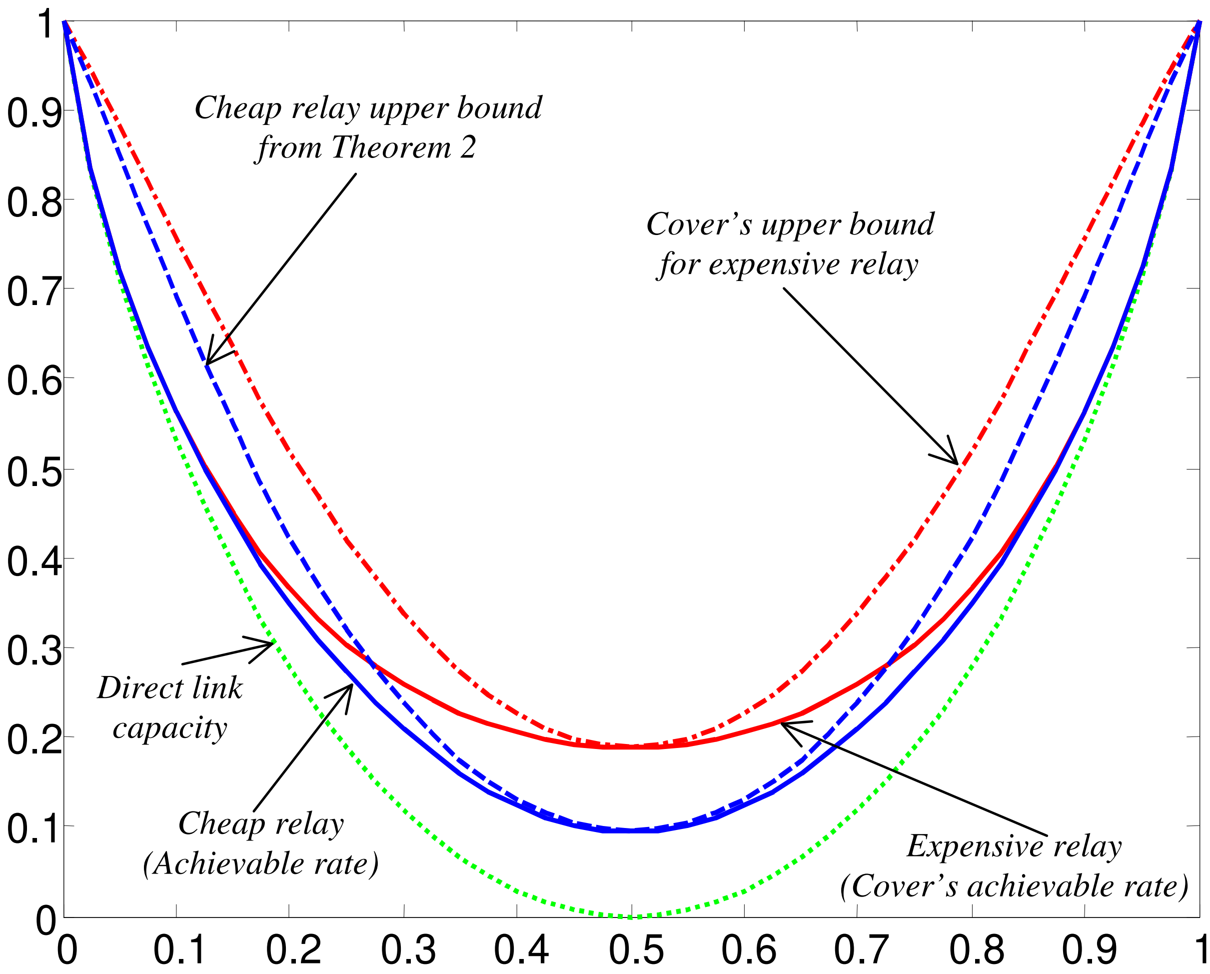

Fig. 4. Inner and outer bounds (in bits/channel use) on the capacity of the cheap relay channel of Figure 2 versus cross over probability of the direct link, p, where each of the links $\mathsf{S} \to \mathsf{D}$, $\mathsf{S} \to \mathsf{R}$, and $\mathsf{R} \to \mathsf{D}$ are binary symmetric channels with cross over probabilities p, q, and r respectively. The q and r take values from the interval $[0, \frac{1}{4}]$ as a function of p, and are symmetric with respect to $p = \frac{1}{2}$, i.e., $q(p) = q(1-p)$ and $r(p) = r(1-p)$.