

Bounds on Achievable Rates for General Multi-terminal Networks with Practical Constraints

Mohammad Ali Khojastepour, Ashutosh Sabharwal, and Behnaam Aazhang

Department of Electrical and Computer Engineering
Rice University
6100 Main St., MS-366
Houston, TX 77005
{amir, ashu, aaz}@rice.edu

Abstract. We consider the problem of communication in a general multi-terminal network where each node of the network is a potential sender or receiver (or both) but it cannot do both functions together. The motivation for this assumption comes from the fact that current radios in sensor nodes operate in TDD mode when the transmitting and receiving frequencies are the same. We label such a radio as a *cheap* radio and the corresponding node of the network as a *cheap* node.

We derive bounds on the achievable rates in a general multi-terminal network with finite number of states. The derived bounds coincide with the known cut-set bound [11] of network information theory if the network has just one state. Also, the bounds trivially hold in the network with cheap nodes because such a network operates in a finite number of states when the number of nodes is finite. As an example, application of these bounds in the multi-hop network and the relay channel with cheap nodes is presented. In both of these cases, the bounds are tight enough to provide converses for the coding theorems [16], and thus their respective capacities are derived.

1 Introduction

Network information theory in general deals with the problem of communication and information transfer between intended subsets of the nodes in a multi-terminal network of senders and receivers. Current results in network information theory are mainly a collection of results for special network topologies or channel models, like: two-way channel, interference channel, multiple access channel, and broadcast channel, and relay channel. The development of multiple user information theory started with the definition of various types of multiple user channels and attempt to find the corresponding capacity regions. Among the investigated channels there are a few channels for which the exact capacity has been derived, and usually the capacity is not known in general.

In this introduction we first give a brief history of the known results in network information theory, which starts with the introduction of various spacial

multiple user channels. Then, we will consider more general results on the multi-terminal network information theory. After that, a short overview of the recent works on characterizing the capacity of wireless network is given, and finally, we summarize the results of this paper.

1.1 Known Results in Network Information Theory

The definition of the two-way channel by Shannon in one of his pioneering papers in 1961 [22] might be considered as the first work on multiple user information theory. It is quite interesting that in different parts of his paper, Shannon also considered two other types of multiple user channels that almost a decade later developed into the multiple access channel and interference channel.

Shannon did not find the capacity result for the general two-way channel (and it is still not known), but in his paper [22], Shannon informally defined a channel (which has later been called as multiple access channel) and mentioned that “a complete and simple solution of the capacity region” for this channel has been found. He never published this result on the multiple access channel, and almost a decade later, in 1971, Ahlswede presented a simple characterization of the multiple access channel [1].

Shannon also considered a ‘restricted’ two-way channel where sender and receiver points at each end are in different places with no direct cross communication. This channel later defined in terms of two pairs of sender-receiver where the transmission of information from each sender to its corresponding receiver interferes with the communications between other sender-receiver pairs [7, 6].

The broadcast channel was also introduced in the same decade. In 1972, Cover first introduced the broadcast channel which involves sending different pieces of information from a single source to several destination [9].

The relay channel was first defined and studied by Van der Mullen [19] for the case of three terminals (or nodes) and later considered by Cover [10]. Cover’s results on the relay channel are still the most comprehensive in terms of achievability and capacity bounds (converse). Although in the relay problem there is just one sender and receiver, it is considered in the domain of network information theory due to the presence of relay node (or nodes).

The aforementioned works were first steps in finding better achievable rates or capacities. Although not many capacity results have been derived, these efforts usually have resulted in derivation of better achievable rates, consideration of special cases and special constraint, investigation of the corresponding channel with continuous alphabet (usually as AWGN channel), and exploration of the effect of feedback.

1.2 Multi-Terminal Network Information Theory

In the aforementioned multiple user channels and in general in network information theory new elements such as interference, cooperation, and feedback make the problem of reliable communications more challenging. Due to the difficulty

of the problem, capacity results are only known for very limited cases with special constraints. In most of the cases where the capacity results are known, they coincide with the cut-set bound on the achievable rates in a network [11]. The bounds have nice and simple max-flow min-cut representation: The rate of flow of information across any boundary is less than or at most equal to the mutual information between the inputs on one side of the boundary (senders' side) and the output on the other side (receivers' side) conditioned on the inputs on receiver's side.

The cut-set theorem is one of the most general results in network information theory. Although the converse (if it exists) to the coding theorem for different channels has been proved independently, in most of the cases it corresponds to a simple representation of the cut-set bounds. For example, the capacity region of multiple access channel is known due to the fact that the discovered achievable region for this channel is large enough to coincide with the cut-set bound. Another example is the relay channel for which an achievable rate was derived by Cover [10]. This achievable rate simply coincides with the cut-set bound for the degraded relay channel, and so the capacity of degraded relay channel is known, while the bound of non-degraded relay channel is higher and is still not shown to be achievable.

On the other hand, there are some cases where an upper bound on the capacity has been derived which is tighter than direct application of the cut-set theorem. For example, the capacity region of general broadcast channel is still not known because either the outer bounds are not tight enough or the achievable rate region is not large enough. But recently, the sum capacity of Gaussian vector broadcast channel (GVBC) has been considered and derived [5, 23, 27, 24], which is lower than the corresponding suggested cut-set bound on the sum rate. The outer bound on this sum rate capacity is derived based on the work of Sato [21] and the achievability part is based on Costa's precoding [8] or Marton's achievable region for the general broadcast channel [18].

Some recent works on network information theory have considered the problem of multi-casting in the networks which can be regarded as the network information flow. Ahlswede et al. [2, 3] introduced a new class of multiuser source coding problems inspired by network communications, and proposed the use of coding in the network for the purpose of multi-casting. This work can be regarded as the max-flow min-cut theorem for the network information flow. Meanwhile, this work reveals that in general coding at the nodes (*network coding*) is required to achieve the best rates and the information should not be treated as a 'fluid' which can simply be routed and replicated at the nodes. Li and Yeung [17] presented a linear code for multi-casting from a single-source which is an example of such a network code. Borade applied the cut-set theorem [11] to the problem of the network information flow and derived an information theoretic upper bound on the information flow in discrete memoryless networks [4].

1.3 Information Theory of the wireless Networks

Since information theoretic analysis of capacity and achievable rates in the network is complicated, it might not give much intuition on construction and limits of practical communication systems. Gupta and Kumar presented another look at the capacity of large networks and characterized the achievable throughput of the nodes [14]. The assumption of point-to-point coding in their work precludes any kind of cooperation and network coding which is possible for example in broadcast or multiple access channels.

New interest in sensor networks and network information theory has sparked new research into the relay channel. Subsequent work by Gupta and Kumar [15] on an extension of the relay channel shows that using more sophisticated multi-user coding schemes can provide sizable gains in terms of *transport capacity*. Also a follow-up paper by Xie and Kumar [26] established an explicit achievable rate expression for the degraded Gaussian channel with multiple relays and characterized the bounds on transport capacity. Reznik et al. [20] considered Gaussian physically degraded relay channels and extended the results for multiple relay stages with a total average power constraint.

Gaspar and Vetterli [12] derived a lower bound on the capacity of the relay channel by allowing arbitrarily complex network coding. Also, they considered upper bounds from the cut-set theorem [11] and showed that these upper and lower bounds meet asymptotically as the number of relay nodes goes to infinity. This result again shows that network coding is essential.

1.4 In This Paper

As we discussed, there are many new elements in network information theory and especially sensor networks that should be considered in the communication problem, such as: interference, feedback, and cooperation. Feedback and cooperation have received most of the attention so far, in order to derive better achievable rates or increase the network capacity in different scenarios, while the effect of interference even in the relay case can be quite capacity limiting and it has not received enough attention. On the other hand another limiting property is the fact that most of the current radios in sensor nodes operate in TDD mode when the transmitting and receiving frequencies are the same. Thus, sensor nodes cannot transmit and receive at the same time, and as a result the achievable rate in the network is lower.

This practical constraint will force us to use the nodes of the network as either a sender or receiver at any given time, which can be interpreted as having more than one mode of operation in the network. Each mode of operation in the network corresponds to the valid partitioning of the nodes into two disjoint subsets of sender and receiver nodes such that there is no node in the sender nodes set which is going to communicate with another node in that subset.

The motivation for the current work comes from the aforementioned practical constraint. We consider a more general case of the problem where the network has a finite number of states and derive some upper bounds on the achievable

information rate in such a network. The derived bounds coincide with the known cut-set bound in network information theory [11] when the network has just one state. On the other hand, the bounds trivially hold for the network with the mentioned practical constraint if nodes are considered as states, and the number of nodes in the network is finite. Later, we will give two applications of the derived bound which provides a tight upper bound where the cut-set bound [11] is unable to provide such bounds. In these two specific examples, we show that the derived bounds are tight enough to characterize the capacity.

Channels with finite number of states have been considered first in the context of compound channels [25], 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 for the AWGN channel with infinite number of states (fading) can also be found in [13]. In the compound channel, besides the fact that knowledge of the channel state information may or may not be available at the sender or receiver, the state is considered to be varying stochastically. Throughout this paper, we assume that the network state information is known to all sender and receiver nodes, and in Corollaries 1, 2 we even let the nodes choose the state of the channel in order to optimize the intended achievable communication rates. As another direct application of the derived bound, Theorem 1 also 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 [25]].

2 Problem Formulation

We consider a general multi-terminal network of senders and receivers. The network can be considered as a directed graph where each node of the network represents a potential sender or receiver (or both), and each link represents existence of a (one-way) channel between two nodes. Most radios operate in TDD mode when the transmitting and receiving frequencies are the same. Thus, they can either send or receive at a given time or use of the network. We label such a radio as a *cheap* radio and the corresponding node of the network as a *cheap* node.

In accordance with the term of *channel use* in a single discrete memoryless channel we define *network use* which corresponds to the simultaneous one time use of each existing nodes of the network either as a sender or receiver. Also in each network use, set of *active links* is defined as all links which are departing from the set of sender nodes. With this definition, the mentioned practical constraint translates to the fact that in each network use there is no active link that arrives at a sender node.

We define operating mode of the network (or state of the network) as a valid partitioning of the nodes of the network into two sets of ‘sender nodes’ and ‘receiver nodes’ such that there is no active link that arrives at a sender node. It is obvious that if the number of nodes in the network is finite then the number of operating modes of the network of cheap nodes are finite.

In the following discussion we consider the network with finite states in general, where the network state is known to all nodes at each time and it is predefined. In other words sequence of the network states is a deterministic function or we can choose it in advance in order to maximize the throughput or optimize other network characteristics.

Results of this section are easily applicable to the network with cheap nodes where operating mode of the network is considered as the state of the network. Thus, although most of the discussion in the rest of paper applies to the network with cheap nodes, it is not limited to such a network. In other words state of the network in this section is not necessarily derived from the mentioned practical constraint.

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

Considering the network of discrete memoryless channels (DMC) with finite number of states and following the notation of [25], the channel transition function or channel probability function (c.p.f.) is defined as: $P(y^{(1)}, y^{(2)}, \dots, y^{(N)} | x^{(1)}, x^{(2)}, \dots, x^{(N)} | m)$ where m is the state of the network. Each network use corresponds to the use of all present channels one time in a specific state. For example in the mentioned network with cheap nodes existing channels in each state is the set of active links or active channels. For every transmitter-receiver node pair (i, j) there is an intended message $W^{(ij)} \in \{1, 2, 3, \dots, 2^{nR^{(ij)}}\}$ to be transmitted from node i to node j in n network uses. The input symbol $X_k^{(i)}$ which is the transmitted signal from node i in time k (or k 'th channel use) depends on $W^{(ij)} \forall j \in \{1, 2, \dots, N\}$, and also the past values of the received signal $Y^{(i)}$, i.e. $Y_1^{(i)}, Y_2^{(i)}, \dots, Y_{k-1}^{(i)}$. Thus, the encoding and decoding functions of block length n code for node i have the following structure and properties:

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

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

The probability of error for each decoder is defined as: $P_e^{(n)(ij)} = Pr(\widehat{W}^{(ij)} \neq W^{(ij)})$ 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 $P_e^{(n)(ij)} \rightarrow 0$ when $n \rightarrow \infty$ for all $j, i \in \{1, 2, \dots, N\}$

Now suppose that state of the network is a deterministic function for every network use k as $m_k, m_k \in \{1, 2, \dots, M\}$, where M is the number of possible states assuming finite number of states. For any state m define $n_m(k)$ as the number of states which is equal to m in the first k network uses. Let $t_m = \lim_{k \rightarrow \infty} n_m(k)/k$ define the portion of the time that network have been used in state m as the total number of network use goes to infinity.

For any cut-set which partitions set of all the nodes into two disjoint set S, S^c we will drive a bound on the information flow from one side to the other side, i.e from sender nodes in the set S to the receiver nodes in the set S^c As Theorem 1 states the bound is not function of the choices of deterministic function m_k directly rather it depends only on the asymptotic values t_1, t_2, \dots, t_M . In fact this theorem bounds the achievable rate of the transmission across a cut-set with the best choice of the deterministic function m_k with the fixed asymptotical properties, i.e. fixed $t_i, i \in \{1, 2, \dots, M\}$

Theorem 1. *Consider a general network with finite number of states, M , for which the sequence m_k of the states of the network is fixed and is known to all nodes. If the information rates $\{R^{(ij)}\}$ are achievable then there exist some joint probability distribution $p(x^{(1)}, x^{(2)}, \dots, x^{(N)}|m)$ such that*

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

for all $S \subset \{1, 2, \dots, N\}$

Proof See appendix I.

Theorem 1 provides a bound on the information flow across any cut-set of the network when the sequence of channel states is a known deterministic function. Thus, considering all possible sequence of Network states the bound across each cut-set would be the supremum of the achievable bounds in Theorem 1 for all choices of deterministic function m_k . Specifically we have the following corollary.

Corollary 1. *Consider a general network with finite states, M , for which the sequence m_k of the states of the network is known to all nodes. Maximum achievable information rates $\{R^{(ij)}\}$ across the cut-set $S \subset \{1, 2, \dots, N\}$ for the proper choice of network state sequence m_k is bounded by:*

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

for some joint probability distribution $p(x^{(1)}, x^{(2)}, \dots, x^{(N)}|m)$ 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 in a fixed state m which allows for the maximum mutual information $I(X_{(m)}^S, Y_{(m)}^{S^c} | X_{(m)}^{S^c})$.

We can also consider an upper bound for the achievable rate of the information flow between two disjoint subsets of the nodes $\{S_1, S_2\}$. This upper bound proves useful when we just try to maximize the rate of information transfer from a given subset of the nodes to the other disjoint subset of the nodes without any interest in other communications in the network. For example this situation happens in the relay problem when a sender node (or set of sender 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 from sender to destination.

A simple observation suggest 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-set S such that $S \subset \{1, 2, \dots, N\}$, $S \cap S_1 = S_1$, $S \cap S_2 = \phi$ which is given by corollary 1. Thus, we have:

$$\sum_{i \in S_1, j \in S_2} R^{(ij)} \leq \min_S \sup_m I(X_{(m)}^S, Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \quad (3)$$

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

We can further elaborate result of theorem 1 directly to find better upper bound for the achievable rate of the information flow between to disjoint subsets of the nodes $\{S_1, S_2\}$.

Theorem 2. *Consider a general network with finite states, M , for which the sequence m_k of the states of the network is fixed and known to all nodes. If the information rates $\{R^{(ij)}\}$ are achievable then there exist some joint probability distribution $p(x^{(1)}, x^{(2)}, \dots, x^{(N)} | m)$ such that sum rate of information transfer from a node set S_1 to a disjoint node set S_2 , $S_1, S_2 \subset \{1, 2, \dots, N\}$, is bounded by:*

$$\sum_{i \in S_1, j \in S_2} R^{(ij)} \leq \min_S \sum_{m=1}^M t_m I(X_{(m)}^S, Y_{(m)}^{S^c} | X_{(m)}^{S^c}) \quad (4)$$

when the minimization is taken over all set $S \subset \{1, 2, \dots, N\}$ subject to $S \cap S_1 = S_1$, $S \cap S_2 = \phi$.

Proof Direct application of the Theorem 1 and considering the fact that the sequence of channel state is fixed results in the mentioned bound. ■

Again, considering all possible sequence of channel states, the bound of the achievable rate of information transfer from a node set S_1 to a disjoint node set S_2 , $S_1, S_2 \subset \{1, 2, \dots, N\}$, is bounded by the supremum of the achievable bounds in theorem 2 for all choices of deterministic function m_k .

Corollary 2. Consider a general network with finite states, M , for which the sequence m_k of the states of the network is known to all nodes. Maximum achievable information rates $\{R^{(ij)}\}$ from a node set S_1 to a disjoint node set S_2 , $S_1, S_2 \subset \{1, 2, \dots, N\}$ for the proper choice of network state sequence m_k is bounded by:

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

for some joint probability distribution $p(x^{(1)}, x^{(2)}, \dots, x^{(N)} | m)$ when the minimization is taken over all set $S \subset \{1, 2, \dots, N\}$ 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.

3 Application of the New Bounds

Theorem 1 provides general bounds on the achievable information rates across any cut-set in the network and these bounds are not necessarily tight. For example, for the networks with just one state of operation, it coincides with the known cut-set bounds [11] which provide tight enough bounds to find the capacity region or capacity rate in some cases like discrete memoryless multiple access channel, degraded discrete memoryless relay channel, or arbitrary discrete memoryless relay channel with feedback [10]. On the other hand, the cut-set bounds [11] for even some other simple examples, such as general (non-degraded) broadcast channel do not provide a tight bound and thus the capacity region is not yet known.

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, but it is not always true. For example in the problem of communication in a network with cheap nodes, 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 obviously just one state of operation but all of the links or channels between the nodes can be used at the same time).

In this section, we will give applications of derived bounds in the previous section in cascaded channels (multi-hop network) and relay channel with cheap nodes. In both of the examples, achievable rate of information transmission is considerably lower than the rate of the same network topology without the mentioned practical constraint. 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 most effective in these cases and they

capture the effect of the introduced practical constraint on the network nodes to the extent that we can actually drive the capacity in both of the examples.

3.1 Cascaded Channels (Multi-hop Network)

Consider L discrete memoryless channel in cascade, Figure 1, and index each channel from left to right as $i = 1, 2, \dots, L$, and each node from left to right as $0, 1, 2, \dots, L$ in which we are interested in transmitting information from node 0 to node L . Thus node i receives Y_i which is the output of the channel i and

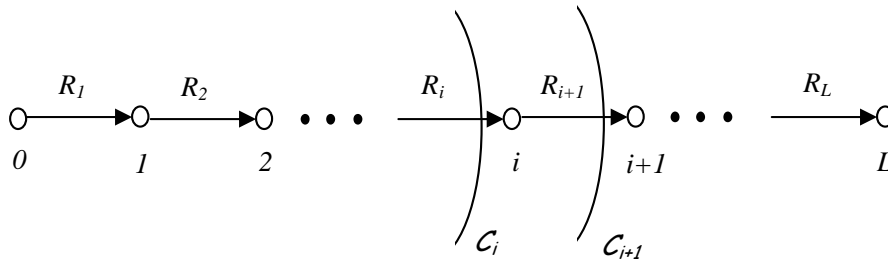


Fig. 1. Cascaded Discrete Memoryless Channels

transmit its information, X_{i+1} via channel $i + 1$ which is the input to this channel. Since we have assumed that channels are cascaded and there are no other connection between the nodes other than stated, the channel output Y_i just depends on the input X_i but no other transmitted signals. For each channel i , $i \in \{1, 2, \dots, L\}$ define capacity of each individual link as $R_i := \max I(X_i, Y_i)$ where maximization is taken over all possible distributions of X_i .

It has been known that the capacity of such cascaded system without the mentioned practical limitation on transmission and reception at the same time is the minimum of the individual rates of the channels $C_1 = \min\{R_1, R_2, \dots, R_L\}$. Since each channel can transmit information at least with the rate of C_1 without any restriction on receiving data from previous node, achievability of this minimum rate is immediate. Also, the known cut-set bound of network information theory states that no higher rate is achievable.

On the other hand, imposing the mentioned practical limitation will decrease the achievable rate in this cascaded channel and the mentioned known cut-set bound is no longer tight. In this case, by using the results of Corollary 2 we have:

$$R \triangleq R^{(0L)} \leq \sup_{t_m} \min_i \left\{ \left(\sum_{m=1}^M t_m * \delta_{im} \right) R_i \right\} \quad (6)$$

when the minimization is taken over i , $i \in \{1, 2, \dots, L\}$ and the supremum is over all the non-negative t_m subject to $\sum_{i=1}^M t_m = 1$. In the above expression $\delta_{im} = 1$ iff link i is used in state m of the network, otherwise $\delta_{im} = 0$.

It is also possible to prove that the above bound is actually achievable. The proof is based on the fact that for any given sets of $\{t_1, t_2, \dots, t_M\}$ associated with the states $1, 2, \dots, M$ satisfying $\sum_{i=1}^M t_m = 1$, the rate of $\min_i \{(\sum_{m=1}^M t_m * \delta_{im}) R_i\}$ is achievable with arbitrarily small probability of error [16]. Thus, the above rate $\sup_{t_m} \min_i \{(\sum_{m=1}^M t_m * \delta_{im}) R_i\}$ is the capacity of the network of the cascaded channels with cheap nodes.

Furthermore, we have proved that the above expression can be simplified to the following form:

$$R = R^* := \min \left\{ \frac{R_1 R_2}{R_1 + R_2}, \frac{R_2 R_3}{R_2 + R_3}, \dots, \frac{R_{L-1} R_L}{R_{L-1} + R_L} \right\} \quad (7)$$

For each $i \in \{1, 2, \dots, L-1\}$, consider two cut-set C_i, C_{i+1} . since for each i, m we have $\delta_{im} * \delta_{(i+1)m} = 0$ it can easily be verified that $R \leq \frac{R_i R_{i+1}}{R_i + R_{i+1}}$ and thus $R \leq R^*$. On the other hand it is possible to show that states $1, 2, \dots, M$ and their associated set of $\{t_1, t_2, \dots, t_M\}$ exist such that the rate R^* is achievable, and thus $R = R^*$ [16].

3.2 Relay Channel

Consider discrete memoryless relay channel of Figure 2, in which source node S is willing to transmit information to the destination node D by using direct link between the node pair (S, D) as well as help of another relay node R (if it improves the achievable rate of transmission) by using link pairs (S, R) and (R, D). Furthermore assume that relay node R is a cheap node and thus it cannot transmit and receive at the same time.

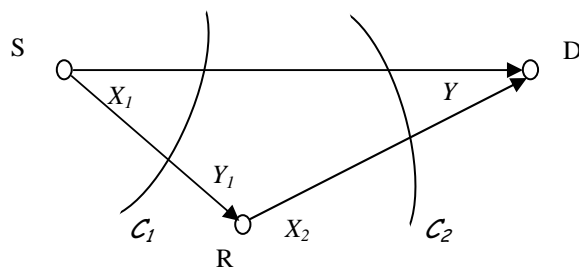


Fig. 2. Discrete Memoryless Relay Channel

With this assumption there are two possible states of operation in the network. In the state m_1 relay node R acts as a receiver and thus channel probability function is given by: $p(y, y_1 | x_1 | m_1)$, while in the state m_2 relay node functions as a transmitter and the channel probability function is given by: $p(y | x_1, x_2 | m_2)$.

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

$$R \leq \sup_{t, 0 \leq t \leq 1} \min \{ t I(X_1; Y, Y_1 | m_1) + (1-t) I(X_1; Y | X_2, m_2), \\ t I(X_1; Y | m_1) + (1-t) I(X_1, X_2; Y | m_2) \} \quad (8)$$

For every t , $0 \leq t \leq 1$, it is possible to show that the rate R^* is achievable where R^* is given by:

$$R^* \triangleq \min \{ t I(X_1; Y_1 | m_1) + (1-t) I(X_1; Y | X_2, m_2), \\ t I(X_1; Y | m_1) + (1-t) I(X_1, X_2; Y | m_2) \} \quad (9)$$

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 the bound of (8) would coincide with this achievable rate for some value of t . Hence, the bound derived in Corollary 2 provides the converse for the capacity theorem of degraded cheap relay channel, and the capacity is given by $C = R^*$ defined in (9).

4 Concluding remarks

Theorems 1, 2, (and Corollary 1, 2) establish general bounds on the achievable information rates in the network with finite number of states. While these bounds are not necessarily tight in general, they provide tight enough bounds for the network of cascaded channels with cheap nodes and also cheap relay network which results in derivation of the capacity in these cases. The derived Bounds coincide with the known cut-set Bounds [11] of network information theory if the network has just one state. In addition, Theorem 1 provides a tight bound for the capacity of the compound channel (which is a multiple state channel) with the known channel state information both at the transmitter and receiver [Theorem 4.6.1 in [25]].

References

1. R. Ahlswede: Multi-way communication channels. in Proc. 2nd Int. Symp. information Theory, Tsahkadsor, Armenian S.S.R., pp. 23-52,, (1971)
2. R. Ahlswede, N. Cai, R.W. Yeung: Network Information Flow Theory, In: Proc. IEEE ISIT 1998, pp. 186, (1998)
3. R. Ahlswede, N. Cai, S.-Y. R. Li, R.W. Yeung: Network Information Flow, IEEE Trans. on Information Theory, Vol.46, No.4, pp. 1204-1216, (2000)
4. S.P. Borade: Network Information Flow: Limits and Acheivability, In: Proc. IEEE ISIT 2002, pp. 139, (2002)
5. G. Caire, S. Shamai: On the Achievable Throughput in Multiple Antenna Gaussian Broadcast Channel, submitted to IEEE Transactions on Information Theory, July (2001)

6. A. B. Carleial: A case where interference does not reduce capacity, IEEE Trans. Inform. Theory, vol. IT-21, pp. 569-570, Sept. (1975)
7. A.B. Carleial: Interference Channels, IEEE Trans. on Information Theory, Vol. IT-24, No. 1, January, pp 60-70 (1978)
8. M. H. M. Costa: Writing on dirty paper, IEEE Transactions on Information Theory, pp. 439 - 441, May (1983)
9. T. M. Cover: Broadcast channels, IEEE Trans. Inform. Theory, vol. IT-18, pp. 2-14, Jan. (1972)
10. Thomas M Cover, Abbas S. El Gammal: Capacity Thorems for the Relay Channel. IEEE Transactions on Information Theory, Vol-25, No.5, pp. 572-584, september. (1979)
11. Thomas M Cover, Joy A Thomas: Elements of Information Theory. John Wiley and Sons, Inc., New York (1991)
12. Michael Gastpar, Martin Vetterli: On the Capacity of Wireless Networks: The Relay Case. (2002)
13. A. J. Goldsmith, P. Varaiya: Capacity of fading channels with channel side information, IEEE Transactions on Information Theory. November (1997)
14. Piyush Gupta and P. R. Kumar: The Capacity of Wireless Networks, IEEE Transactions on Information Theory, pp. 388-404, vol. IT-46, no. 2, March (2000)
15. P.Gupta, P.R. Kumar: Towards an Information Theory of Large Networks: An Achievable Rate Region, Submitted to the IEEE Trans. on Information Theory, September (2001), Revised November (2002)
16. Mohammad A. Khojastepour, Ashutosh Sabharwal, Behnaam Aazhang: On the Capacity of 'Cheap' Relay Networks. to be submitted to the 37th Annual Conference on Information Sciences and Systems, Baltimore, Maryland, March 12-14 (2003)
17. S.-Y. R. Li, R.W. Yeung: Linear Codes for Network Information Flow, In: Proc. IEEE ISIT 1999, pp. 25, (1999)
18. K. Marton: A coding theorem for the discrete memoryless broadcast channel, IEEE Transactions on Information Theory, pp. 306 - 311, May (1979)
19. E. C. van der Meulen: Three-terminal communication channels, In: Adv. Appl. Prob., vol. 3, pp. 120-154, (1971)
20. A. Reznik, S.R. Kulkarni, S. Verdu: Capacity and Optimal Resource Allocation in the Degraded Gaussian Relay Channel with Multiple Relays, In: 40th Allerton Conference on Communication, Control, and Computing, September (2002)
21. H. Sato: An outer bound to the capacity region of broadcast channels, IEEE Trans. on Information Theory, Vol. 24, No. 3, pp. 374-377, May (1978)
22. C. E. Shannon: Two-way communication channels, In: Proc. 4th Berkelev Svm . Math. Statist. and Prob.. vol. 1. VD. 611-644. (1961). Reprinted in: Key Papers in the Development-of Information Theory, (D. Slepian, Ed.) New York: IEEE Press, pp. 339- 372, (1974)
23. S. Vishwanath, N. Jindal, and A. Goldsmith: On the Capacity of Multiple Input Multiple Output Broadcast Channels, International Conference on Communications, (2002)
24. Pramod Viswanath, David Tse: Sum Capacity of the Multiple Antenna Gaussian Broadcast Channel and Uplink-Downlink Duality , submitted to IEEE Transactions on Information Theory, July (2002)
25. Jacob Wolfowitz : Coding Theorems of Information Theory. 3rd edn. Springer-Verlog, Berlin Heidelberg Newyork (1978)
26. L. Xie, P.R. Kumar: A network information Theory for Wireless Communication: Scaling Laws and Optimal Operation, Submitted to the IEEE Trans. on Information Theory, April (2002)

27. W. Yu, and J. Cioffi: Sum Capacity of Gaussian Vector Broadcast Channels, submitted to IEEE Transactions on Information Theory, Nov., (2001)

Appendix I: Proof of Theorem 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. Based on the definition of $W^{(ij)}$ such that they are independent and uniformly distributed over their range and also following the definitions and conditions of Section 2, we have:

$$\begin{aligned} & \sum_{i \in S, j \in S^c} R^{(ij)} \\ &= \sum_{i \in S, j \in S^c} H(W^{(ij)}) \end{aligned} \quad (10)$$

$$= H(W^T) \quad (11)$$

$$= H(W^T | W^{T^c}) \quad (12)$$

$$\begin{aligned} &= I(W^T, Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_N^{(S^c)} | W^{(T^c)}) \\ &\quad + H(W^T | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_N^{(S^c)}, W^{(T^c)}) \end{aligned} \quad (13)$$

$$\leq I(W^T, Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_N^{(S^c)} | W^{(T^c)}) + n\epsilon_n \quad (14)$$

$$= \sum_{k=1}^N I(W^T, Y_k^{(S^c)} | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_{k-1}^{(S^c)}, W^{(T^c)}) + n\epsilon_n \quad (15)$$

$$\begin{aligned} &= \sum_{k=1}^N \{H(Y_k^{(S^c)} | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_{k-1}^{(S^c)}, W^{(T^c)}) \\ &\quad - H(Y_k^{(S^c)} | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_{k-1}^{(S^c)}, W^{(T^c)}, W^T)\} + n\epsilon_n \end{aligned} \quad (16)$$

$$\begin{aligned} &\leq \sum_{k=1}^N \{H(Y_k^{(S^c)} | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_{k-1}^{(S^c)}, W^{(T^c)}, X_k^{(S^c)}) \\ &\quad - H(Y_k^{(S^c)} | Y_1^{(S^c)}, Y_2^{(S^c)}, \dots, Y_{k-1}^{(S^c)}, W^{(T^c)}, W^T, X_k^{(S^c)}, X_k^{(S)})\} + n\epsilon_n \end{aligned} \quad (17)$$

$$\leq \sum_{k=1}^N \{H(Y_k^{(S^c)} | X_k^{(S^c)}) - H(Y_k^{(S^c)} | X_k^{(S^c)}, X_k^{(S)})\} + n\epsilon_n \quad (18)$$

$$= \sum_{k=1}^N I(X_k^{(S)}, Y_k^{(S^c)} | X_k^{(S^c)}) + n\epsilon_n \quad (19)$$

$$= \sum_{k=1}^N I(X_{Q_m(k)}^{(S)}, Y_{Q_m(k)}^{(S^c)} | X_{Q_m(k)}^{(S^c)}, Q_m(k) = k) + n\epsilon_n \quad (20)$$

$$= \sum_{m=1}^M n_m(n) I(X_{Q_m}^{(S)}, Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}, Q_m) + n\epsilon_n \quad (21)$$

$$= \sum_{m=1}^M n_m(n) \{H(Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}, Q_m) - H(Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}, X_{Q_m}^{(S)}, Q_m)\} + n\epsilon_n \quad (22)$$

$$\leq \sum_{m=1}^M n_m(n) \{H(Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}) - H(Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}, X_{Q_m}^{(S)})\} + n\epsilon_n \quad (23)$$

$$= \sum_{m=1}^M n_m(n) I(X_{Q_m}^{(S)}, Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}) + n\epsilon_n \quad (24)$$

where

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

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

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

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

(14) 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 $\{R^{(ij)}\}$ are achievable; and also note that $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$),

(15) follows from the chain rule,

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

(17) 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,

(18) 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),

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

(20) follows from introducing time sharing random variables Q_1, Q_2, \dots, 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 , $k \in \{1, 2, 3, \dots, N\}$, for which $m_k = m$,

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

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

(23) 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 ,

(24) follows from the definition of the mutual information.

Thus, by dividing both sides of inequality (24) by n and finding the limit as $n \rightarrow \infty$ we have:

$$\sum_{i \in S, j \in S^{(c)}} R^{(ij)} \leq \sum_{m=1}^M \frac{n_m(n)}{n} I(X_{Q_m}^{(S)}, Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}) + \epsilon_n \quad (25)$$

and as $n \rightarrow \infty$ we have:

$$\sum_{i \in S, j \in S^{(c)}} R^{(ij)} \leq \sum_{m=1}^M t_m I(X_{Q_m}^{(S)}, Y_{Q_m}^{(S^c)} | X_{Q_m}^{(S^c)}) \quad (26)$$

and it completes the proof. ■