On Achievable Information Rates in Single-Source Non-Uniform Demand Networks

Chandra Chekuri
Lucent Bell Labs
Murray Hill NJ 07974, USA
Email: chekuri@lucent.com

Christina Fragouli
School of Comput. and Commun. Sciences
EPFL, Lausanne
Email: christina.fragouli@epfl.ch

Emina Soljanin
Lucent Bell Labs
Murray Hill NJ 07974, USA
Email: emina@lucent.com

***Abstract*— A non-uniform demand network consists of a source and a set of receivers that have different min-cut values from the source. We look at the case where each receiver would like to receive information from the source at a rate that is equal to its min-cut value. This problem has been formulated before, and in constrast to the uniform case, it has been shown that the non-uniform case does not admit a good characterization. Motivated by this, we formulate relaxations of the problem and present some preliminary results.**

I. Introduction

Network coding introduced in [1], [2] considered the multicast setting in which all receivers had an identical min-cut value from a source. In [3], network coding was considered for the non-uniform setting where receivers can have different min-cut values from the source, and it was shown that the non-uniform setting does not admit a nice characterization. In this paper, we formulate and study generalizations of the non-uniform problem when general information transmission mechanisms including routing, network coding, and time sharing are allowed. We first introduce a problem formulation and then study the space of achievable rates in the formulation and point to some interesting phenomena.

1) Problem Formulation: Consider a network represented by a graph $G = (V, E)$, an information source S located on a node of the graph that produces bits at rate ω, and k disjoint groups $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_k$ of receivers. We can equivalently think of the source S to consist of ω unit rate information sources $S_1, \ldots, S_\omega$ located at the same node of G. We assume that all edges in E have unit capacity, and that the min-cut value from the root to each of the $N_i = |\mathcal{R}_i|$ receivers in group $\mathcal{R}_i$ is h_i. Without loss of generality, we assume that $\omega \geq h_1 > h_2 > \ldots > h_k$. Each receiver would like to receive information at a rate equal to its min-cut value. We call this configuration a *non-uniform demand* network.

Definition 1: We say that a pair (α, β) with $\alpha, \beta \in [0, 1]$ is *achievable* for a given instance if there is a transmission mechanism (network and channel coding, routing, time sharing) such that, for $1 \leq i \leq k$, some βN_i of the receivers in $\mathcal{R}_i$ receive rate αh_i.

Our goal is to understand and characterize the space of achievable pairs (α, β). For $k = 1$ and for directed graphs, the pair $(1, 1)$ is clearly achievable by network coding [1], [2]. Here we will be interested in the case where $k > 1$.

2) Motivation and Related Work: The motivation for considering this problem is two-fold, one practical and the other theoretical. Consider an Internet application where a source is transmitting video simultaneously to different receivers. Non-uniform min-cut values arise when the receivers (because of network topology or pricing) have connections with different bandwidths. In this setting, a receiver expects the quality of its stream to be proportional to its connectivity (min-cut value, purchased bandwidth), and this can be accomplished by the source by using multiple description coding. The problem we consider models this application.

Second, we believe that the problem is interesting from a theoretical point of view. The seminal result of Ahlswede et al. [1] completely solved the multicast problem by a simple characterization. However, in more general scenarios with multiple sources and receivers where each receiver has a specific source (or set of sources) from which it requires information, it is much harder to characterize solvability. It has been established that natural necessary conditions based on min-cut values are not sufficient for solvability, see for example [3], [4], [5]. In addition, the problem of deciding if an instance is solvable is often intractable (NP-complete or worse). Therefore, it is of interest to consider simpler situations where all the information sources are located at the same node in the network. Even with this restriction, we can still distinguish two cases. The first is when each receiver specifies the set of sources that it wants information from. The second case, which is the focus of this paper, is when the receivers care only about the rate of information that they receive from the source node. Note that this problem is of interest only if the receivers have different min-cut values from the source, for otherwise we know that the problem is solvable. Even in this case, it is non-trivial to understand the tradeoffs between routing, network coding, time sharing and other transmission mechanisms and we believe it is of interest to study the problem in depth.

The problem we consider has been first formulated and studied in [3], where it was examined whether each receiver can obtain rate equal to its min-cut from the source. In our notation, this is the same as asking if the point $(1, 1)$ is achievable in all networks. It was established in [3] that $(1, 1)$ is not feasible in all instances and further, the problem of deciding whether $(1, 1)$ is feasible is NP-complete. The negative results in [3] motivated us to relax the problem by introducing parameters α and β as already described. From [3] it follows that, given an instance, it is NP-hard to find the largest α such that $(\alpha, 1)$ is achievable. Similarly, it is also NP-hard to find the largest β such that $(1, \beta)$ is achievable.

3) Organization: In Sec. II, we describe two classes of network configurations on which several examples in the paper are based. In Sec. III, we present a family of directed graphs for which $(1, \beta)$ is achievable only for $\beta = O(1/N)$, where $N = \sum_i N_i$ is the total number of receivers. In Sec. IV, we present simple schemes for directed graphs that achieve $(\alpha, 1)$, with α depending on the network parameters. We also show that in undirected graphs $(1/2, 1)$ is achievable using only routing. In Sec. V, we discuss a number of interesting issues that arise in the non-uniform demand scenario but do not play a role in the uniform demand case.

II. Network Topologies

1) Combination Networks: A class of networks known as combination networks is illustrated in Fig. 1. There are three

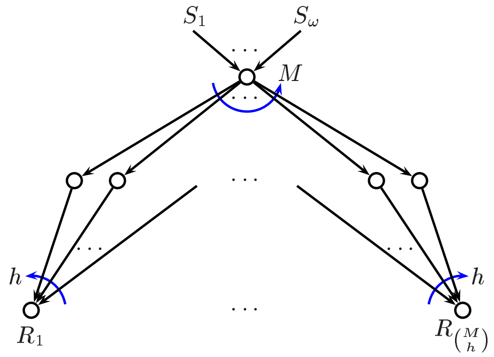

Fig. 1. Combination $B(h, M)$ network.

layers of nodes. The first layer contains the source node, at which ω information sources are located (typically $\omega = h$). The second layer contains M nodes connected to the source node. The third layer contains $\binom{M}{h}$ receiver nodes. Each receiver corresponds to a distinct subset of h nodes from the second layer, and is connected to those h nodes.

2) ZK *Networks:* The ZK networks are a class of networks introduced in [6] in connection with a combinatorial optimization problem called the directed Steiner tree problem. Let N and p, $p \leq N$, be two integers and $\mathcal{I} = \{1, 2, \ldots, N\}$ be an index set. We define two more index sets: $\mathcal{A}$ as the set of all $(p-1)$-element subsets of $\mathcal{I}$ and $\mathcal{B}$ as the set of all p-element subsets of $\mathcal{I}$. A $\text{ZK}(N, p)$ network, illustrated in Fig. 2, is defined by two parameters N and p as follows: node S has w co-located sources each transmitting a unit rate to N receiver nodes $R_1 \ldots R_N$ through a network of three sets of nodes A, B and C. In the typical case we have $w = h = \binom{N-1}{p-1}$ with h as the min-cut from S to each receiver. A-nodes are indexed by the elements of $\mathcal{A}$, and B and C-nodes, by the elements of $\mathcal{B}$. An A node is connected to a B node if the index of A is a subset of the index of B. A B node is connected to a C node if and only if they correspond to the same element

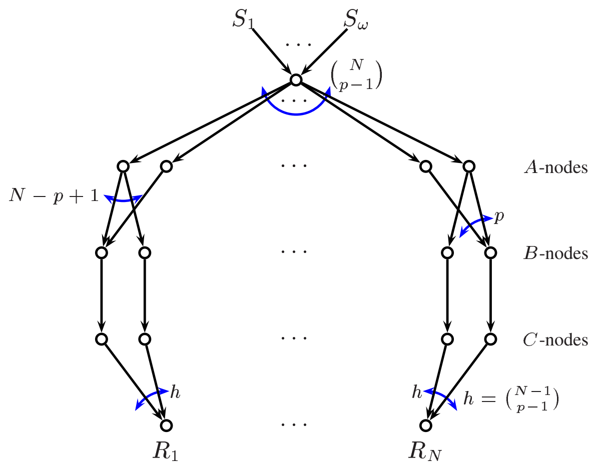

Fig. 2. $\text{ZK}(N, p)$ network.

of $\mathcal{B}$. A receiver node is connected to the C nodes whose indices contain the index of the receiver. ZK networks were used in [7] as examples to show that network coding is an essential transmission mechanism for the receivers to receive a large fraction of their min-cut rate. We use this insight to construct non-unifrom demand networks using ZK networks as a building block.

III. Bounds on β, for $\alpha = 1$

Proposition 1: There are networks with $2N'$ receivers for which no point $(1, \beta)$ where $\beta > 2/N'$ is achievable.

Proof: Starting with the $\text{ZK}(N, p)$ network, we construct a new network as follows. We remove all the A-nodes whose labels contain N and all edges incident to the removed nodes [7]. To each of the $\binom{N-1}{p}$ C-nodes whose corresponding B nodes have multiple inputs, we connect a receiver through a unit capacity edge. We now have $N' = N - 1$ receivers in group $\mathcal{R}_1$ with min-cut $h_1 = \binom{N-1}{p-1}$ and $\binom{N-1}{p}$ receivers in group $\mathcal{R}_2$ with min-cut $h_2 = 1$.

Consider one of the receivers in $\mathcal{R}_2$ which achieves the rate equal to its min-cut. This receiver must be connected to a C node whose corresponding multi-input B node only forwards one of its inputs. Now, there are p receivers in $\mathcal{R}_1$ connected to this C–B node pair, and $p-1$ of them will receive the rate of at most $h_1 - 1$ [7]. Therefore,

$$\beta \leq \frac{N - p + 1}{N}.$$

The claim follows by setting $p = N' - 1$. ∎

We next illustrate a tradeoff between α and β:

Proposition 2: If the point $(1, \beta)$ is achievable for all instances, so are the points $(1/m, 1 - (1 - \beta)^m)$ over $m \geq 1$ time slots.

Proof: During the first time-slot, a fraction β of the receivers observes their min-cut rate. During the second time-slot, an additional fraction $\beta(1-\beta)$ observe their min-cut rate (fraction β of the remaining $1-\beta$ receivers), resulting in

$$\beta + \beta(1-\beta) + \beta(1-\beta)^2 + \ldots \beta(1-\beta)^{m-1} = 1 - (1-\beta)^m$$

receivers observing $1/m$ of their respective min-cut rates. ∎

IV. Bounds on α, for $\beta = 1$

We derive several simple lower bounds on α for directed graphs and a more complex one for undirected graphs.

Proposition 3: In directed graph instances, points $(\alpha, 1)$ are achievable for the following values or ranges of α:

1) $\alpha \geq h_k/h_1$,
2) $\alpha = 1/k$,
3) $\alpha = 1/(k - \sum_{i=1}^{k-1} h_{i+1}/h_i) \geq 1/(m \log_m h_1/h_k)$ for $m > 1$.

Proof:

1) Pretend that all receivers have min-cut h_k (the smallest value) and then apply network coding.
2) Use time sharing and network coding. Consider k time-slots, and allocate one to each set of receivers. Perform network coding for group $\mathcal{R}_i$ during time-slot i.
3) This refines the idea of the previous claim. Consider L time-slots and partition them into k sets so that the size of the ith set is l_i and $l_k + l_{k-1} + \ldots + l_1 = L$. The transmission mechanism is a mix of time sharing and network coding. For $k \geq i \geq 1$, during the l_i time slots, information is transmitted to all receivers in $\cup_{j \leq i} \mathcal{R}_j$ using network coding at a rate of h_i per time slot. Note that this is possible since all receivers in $\cup_{j \leq i} \mathcal{R}_j$ have min-cut at least h_i. Thus a receiver in $\mathcal{R}_i$ receives information in $l_k + l_{k-1} + \ldots + l_i$ time slots albeit at different rates. Simple algebraic manipulation shows that this scheme ensures $\alpha = 1/(k - \sum_i h_{i+1}/h_i)$ using $l_k = \alpha L$ and $l_i = \alpha L(1 - h_{i+1}/h_i)$ for $i < k$. It can be verified that $k - \sum_i h_{i+1}/h_i \leq m \log_m h_1/h_k$ for $m \geq 2$. We mention this since a bound of $\alpha = 1/(m \log_m h_1/h_k)$ is obtained via a grouping of receivers using geometric intervals of the form $[m^i, m^{i+1})$.

∎

Remarks: Although time sharing schemes suggested by Claims 2) and 3) seem simplistic, there exist directed graph instances for which they are close to the best possible and much more powerful than plain routing. Consider the ZK$(N, \sqrt{N})$ family of graphs with two sets of receivers as described in the proof of Prop. 1. Claim 2) guarantees $\alpha = 1/2$. On the other hand, if we use only routing, it is shown in [7] that $\alpha = O(\frac{1}{\sqrt{N}})$ for the receivers with min-cut h_1. If network coding is used to achieve a high rate for receivers in $\mathcal{R}_1$, some receiver in $\mathcal{R}_2$ essentially obtains no information which results in low α. This example demonstrates the effectiveness of time sharing for receivers that differ substantially in their min-cut values.

We next show that, in *undirected graphs*, routing alone is sufficient to achieve $\alpha \geq 1/2$. It is interesting to consider this result in the light of an earlier result for the $k = 1$ case [8], where it was shown that network coding can at most double the achievable throughput in undirected networks. Our result shows that, for nonuniform demand networks and $k > 1$, the benefits of network coding over routing in undirected networks can also be at most a factor of two.

Theorem 1: In undirected graph instances, $\alpha \geq 1/2$ is achievable even if only routing is allowed.

To prove this result, we need some background material. As has been noted in prior work, the information transmission capacity if only routing is allowed, is intimately related to the problem of packing Steiner trees in a graph [8], [9], [10], [7]. In a directed graph, a Steiner tree for a root node $r \in V$ and a set of *terminals* R is an out-tree T rooted at r such that there is a path in T from r to each terminal in R. A *partial* Steiner tree is one in which only some of the terminals in R can be reached from r. Let $R(T)$ denote the reachable set of terminals in T. Given a partial Steiner tree T, we say that a terminal $v \in R$ occurs in T if $v \in R(T)$. For a vertex $v \in G$ let $\lambda(v)$ denote its min-cut value from r. Edmonds [11] showed that if $\lambda(v) \geq k$ for all $v \in V$, then there are k edge-disjoint spanning trees in G. Such a decomposition result does not hold for the Steiner case. That is, if we have $\lambda(v) \geq k$ for all $v \in R$, then the maximum number of edge disjoint Steiner trees on R could be as small as $O(k/\sqrt{|R|})$ [6]. However if the directed graph satisfies a strong requirement, then Bang-Jensen, Frank and Jackson [12] obtained the following result. To state this we let $\rho(v)$ and $\delta(v)$ denote the in-degree and out-degree of v in G.

Theorem 2 ([12]): Let $G = (V, E)$ be a directed graph, $r \in V, R \subset V$ such that for all $v \in V \setminus (\{r\} \cup R)$, we have $\rho(v) \geq \delta(v)$. Let $k = \max_{u \in R} \lambda(u)$. Then there are k edge-disjoint partial Steiner trees $T_1, T_2, \ldots, T_k$ such that each node $u \in R$ occurs in $\lambda(u)$ of the k trees.

We note that [12] also gives a strongly polynomial time algorithm to compute the above decomposition. A strongly polynomial time algorithm requires an implicit representation: identical trees are represented via a single copy and an integer multiplicity value.

Corollary 1: Let G be an undirected graph. Let $k = \max_{u \in R} \lambda(v)$. Then there are k partial Steiner trees $T_1, T_2, \ldots, T_k$ such that each node $u \in R$ occurs in $\lambda(u)$ of the k trees. Further each edge e participates in at most two of the partial trees.

Proof: We obtain a directed graph from an undirected graph by taking each undirected edge e and replacing it by two oppositely directed edges e_1 and e_2. It is easily seen that for any vertex v, the min-cut value in the resulting directed graph is no less than the min-cut value in the original undirected graph. We then apply Theorem 2. ∎

Now we are ready to prove Theorem 1.

Proof: We apply Corollary 1 to our given graph G and the set of receivers $\mathcal{R} = \cup_i \mathcal{R}_i$. Let $T_1, T_2, \ldots, T_{h_1}$ be the resulting trees. A receiver $u \in \mathcal{R}_i$ occurs in h_i of the trees. Note that edge e can be in two different trees, used in each tree in opposite directions. Therefore the source can independently transmit at rate $1/2$ in each of these trees while obeying capacity. A receiver $u \in \mathcal{R}_i$, receives a rate of $h_i/2$ since it occurs in h_i distinct trees. ∎

Theorem 1 shows that there are no significant throughput benefits to be expected from use of network coding in undirected networks. The benefits of network coding in such a scenario would be in terms of complexity, when simple strategies, like the ones described in Prop. 3, would be effective. That is, in networks with a small number of groups of receivers, or where the min-cuts of the different receivers groups do not differ significantly, we can use the simple network coding strategies to achieve with very low complexity a constant fraction of the min-cut.

V. Some New Issues

In the usual multicast setting, it is assumed that there are h sources and that each receiver has min-cut h from the sources. Here we describe some interesting issues that arise if we deviate from this particular setting even slightly. We highlight the significance of source co-location, the number of transmission time slots, and the use of memory for the $(\alpha, 1)$ case.

A. Source Co-location

As mentioned before, we can think of our single source as ω co-located sources, without altering the problem formulation. However, if the ω sources are located in different nodes of the graph, even if the receivers still only want to receive rate equal to their min-cut promiscuously from all sources, the problem changes quite significantly. Such a situation might arise in peer-to-peer networks content distribution.

Example 1: Consider the graph with three sources and three receivers shown in Fig. 3. Note that all receivers have min-cut two, but it is not possible for all receivers to get rate equal to two simultaneously. On the contrary, if the three sources were co-located, we could multicast to each receiver at rate two.

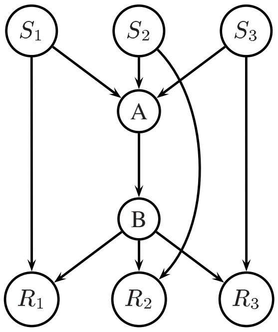

Fig. 3. A configuration with three sources and three receivers.

We generalize Example 1 and discuss a configuration where, if time sharing is not allowed, the ratio between the achievable α with co-located sources and separate sources can be made $1/h$.

Example 2: Consider a combination network with ω sources that are not co-located. There exist $N = \omega\binom{M}{h-1}$ receivers, where each receiver directly observes one of the sources, and a subset of size $h-1$ of the M nodes. In Example 1 $\omega = 3$, $M = 1$, and $N = 3$. Each of the M nodes will receive and forward a linear combination that will depend on x of the ω sources, with $1 \leq x \leq h$. For M large enough, from the pigeonhole principle, there will exist receivers that will experience throughput one, either because they observe h times the same source, or because the $h-1$ linear equations they receive through the M nodes have too many unknowns. Note that for $\omega = h$ network coding can be used to send rate h to each receiver.

B. The Number of Time Slots

Our observation here is that, if we are restricted to convey the information within L time-slots, and L is bounded, we can construct networks large enough so that for $\beta = 1$, α takes small values. However, if we let L grow with the network size, simple techniques achieve a constant α.

The following example describes a non-uniform demand network where using large number of time slots, $1 - 1/e$ is achievable. We believe that if L is a fixed constant and we choose the network sufficiently large then α goes to zero.

Example 3: Consider a combination network with h sources and M nodes in the second layer. We have k receiver groups, where group $\mathcal{R}_i$ has min-cut

$$h_i = 2^{r(k-i+1)} \quad i = 1 \dots k, \quad r > 1$$

and number of receivers

$$N_i = \binom{M}{h_i}.$$

That is, each receiver in $\mathcal{R}_i$ is connected to a distinct subset of M-nodes. Note that we can choose the parameter k to be much larger than L, so that the simple techniques in Prop. 3 either cannot be used, or result in an α that is arbitrarily small.

Moreover, for fixed L, we believe we can make the network large enough so that even if we perform network coding, there will exist at least one receiver that after L time-slots will have a received a fraction $\alpha \to 0$ of its min-cut value.

On the contrary, if we allow the number of time-slots to grow, while the graph is fixed, we can perform randomized routing as described in [7] and achieve a fraction $\alpha \geq (1-1/e)$ for all receivers.

C. Memory

We discuss how the achievable common rate α can be substantially impacted by allowing intermediate nodes to store information that they receive and forward it selectively in the future time slots. Example 4 illustrates this. The basic idea is that, in a network with multiple layers, if a node receives some information and cannot immediately convey all of it to its descendent layers, then it can store it, and possibly send it to the interested receivers at a later time slot.

Example 4: We construct an example configuration, where, if intermediate nodes are allowed to have memory, $\alpha \geq 1/2$ is achievable with a simple scheme, while otherwise α may go to zero.

The configuration consists of a gadget network, that is repeated to create a multi–layer network. The gadget is a variation of the ZK(N,p) network. It has the sets of nodes A, B and C as described in Section II. It additionally has two new sets of nodes, E and F.

Each node in E is connected to a different node in C through a unit capacity edge. Thus the min-cut to nodes in E is one. Recall that each node in C is associated with a p-element subset. Thus we can think of each node in E to also be associated with a p-element subset, the same as the node C it is connected with. Each node in F is connected to $\frac{N}{p}$ distinct E nodes[1], corresponding to mutually disjoint subsets of size p. Note that each node in F, is not connected to any node in set A through more than one paths. As a result, each node in F observes the information flowing through a given node A at most once. The gadget contains two sets of receivers.

- *Set* $\mathcal{R}_1$: Contains the usual N receivers of the ZK(N,p) network, with min-cut $\binom{N-1}{p-1}$.
- *Set* $\mathcal{R}_2$: Contains $\binom{N}{p}$ receivers, one on each of the nodes in set E, with min-cut 1.

This graph will form our basic gadget. We will call this new network $G_Z(N,p)$. We create the concatenated network as follows. The first level consists of one copy of $G_Z(N_1,p_1)$. To create the second level, we attach a new such gadget network $G_Z(N_2,p_2)$ to each node in F, with

$$\frac{N_1}{p_1} = \binom{N_2-1}{p_2-1}.$$

We let $\mathcal{R}_1^i$ and $\mathcal{R}_2^i$ denote the receivers in level i of the construction and these correspond to $\mathcal{R}_1$ and $\mathcal{R}_2$ in the basic gadget $G_Z(N,p)$. Also let F^i denote the set of all F-nodes in level i. In the second level of the network we will have receivers that belong to group $\mathcal{R}_2^2$ with min-cut 1, and receivers in the group $\mathcal{R}_2^2$ with min-cut $\frac{N_2}{p_2}$. We can continue recursively adding r levels.

Assume that we allow intermediate nodes in the network to use memory. We can achieve $\alpha = 1/2$, using the following transmission strategy that uses time sharing, routing, network coding, and memory. Consider two time slots T_1 and T_2. During T_1 we transmit information using simple routing so that all receivers in $\cup_i \mathcal{R}_2^i$ and all nodes in $\cup_i F^i$ receive information at rate equal to their min-cut value. Consider the first level $G_Z(N_1,p_1)$. The source sends distinct bits to nodes in level A. Each node in level B picks an arbitrary node in level A to receive a bit, and forwards it to the corresponding E node which in turn forwards it to the F node it is connected to. From the construction of the network, it can be verified that each F node gets distinct bits of information and hence receives its min-cut rate. Now, each F node acts as a source for second level and a similar transmission mechanism can be used. Thus, in T_1, all nodes in $\cup_i \mathcal{R}_2^i$ receive information at their min-cut rate.

During T_2 no information is conveyed from layer to layer. Instead, all source nodes $(\cup_i F^i)$ use their stored bits to perform network coding for the receivers in groups $\cup_i \mathcal{R}_1^i$. Thus all receivers in groups $\mathcal{R}_1^i$ receive their full rate in T_2. This shows that $\alpha = 1/2$ is achievable.

Assume now that the intermediate nodes in the network do not have memory. It is easy to see that the simple timesharing strategy fails. This is because of the following properties.

- If we perform routing for $\cup_i \mathcal{R}_2^i$, receivers in group $\mathcal{R}_1^i$ will not observe a constant fraction of their min-cut [7].
- If we perform network coding for group $\mathcal{R}_1^i$, the receivers in groups $\mathcal{R}_2^i$, as well as the nodes F will observe more equations than unknowns, and thus will not be able to decode. That is, no rate will be conveyed from layer to layer.

It is of course possible to think of elaborate and involved schemes that combine routing, coding, and time sharing. We believe that for the construction we describe, a constant α is not achievable without using memory. Although we are unable to formally prove this for all schemes, we can show the unachievability for some restricted but more elaborate schemes than simple time sharing.

We observe that receivers in $\mathcal{R}_1^i$ for different values of i have substantially different min-cut values. Thus, any of the simple schemes described in Prop. 3 would not achieve constant α. The ZK network has been picked as the basic gadget since network coding is essential for achieving a good rate to $\mathcal{R}_1$, as established in [7]. The min-cut to the F nodes that act as sources to the next level, is substantially smaller than the min-cut to the receivers in $\mathcal{R}_1$, and we hypothesize that time sharing is the only way to achieve constant rate for both the F nodes and the $\mathcal{R}_1$ in the basic gadget $G_Z(N,p)$. We believe that by modifying this example, it is possible to show that a constant α is not achievable in directed graphs even if memory is allowed.

[1] For simplicity we assume that these fractions are integers, for example by selecting $p_i = \sqrt{N_i}$.

REFERENCES

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

[2] S-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, pp. 371–381, Feb. 2003.

[3] Y. Cassuto and J. Bruck, "Network coding for nonuniform demands," *ISIT*, 2005.

[4] A. Lehman, *Network Coding*. PhD thesis, MIT, 2005.

[5] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *Submitted to Trans. Inf. Theory*, 2004.

[6] L. Zosin and S. Khuller, "On directed steiner trees," in *Proc. 13th Annual ACM/SIAM Symposium on Discrete Algorithms (SODA '02)*, pp. 59–63, 2002.

[7] C. Chekuri, C. Fragouli, and E. Soljanin, "On average throughput and alphabet size in network coding," *To appear in IEEE Trans. Inform. Theory*, June 2006.

[8] Z. Li and B. Li, "Network coding in undirected networks," *Proceeding of CISS 2004*, 2004.

[9] Y. Wu, P. A. Chou, and K. Jain, "A comparison of network coding and tree packing," *ISIT 2004*, 2004.

[10] A. Agarwal and M. Charikar, "On the advantage of network coding for improving network throughput," in *Proc. of 2004 IEEE Information Theory Workshop*, San Antonio, Texas, Oct. 2004.

[11] J. Edmonds, "Edge-disjoint branchings," *Combinatorial Algorithms (B. Rustin, ed.), Acad. Press, New York*, pp. 91–96, 1973.

[12] Bang-Jensen and A. Frank and B. Jackson, "Preserving and increasing local edge-connectivity in mixed graphs," *SIAM J. on Discrete Math*, vol. 8, pp. 155–178, 1995.