# Network Coding: The Case of Multiple Unicast Sessions

Zongpeng Li, Baochun Li
Department of Electrical and Computer Engineering
University of Toronto
{arcane, bli}@eecg.toronto.edu

**Abstract**

In this paper, we investigate the benefit of network coding over routing for multiple independent unicast transmissions. We compare the maximum achievable throughput with network coding and that with routing only. We show that the result depends crucially on the network model. In directed networks, or in undirected networks with integral routing requirement, network coding may outperform routing. In undirected networks with fractional routing, we show that the potential for network coding to increase achievable throughput is equivalent to the potential of network coding to increase bandwidth efficiency, both of which we conjecture to be non-existent.

## 1 Introduction

We study in this paper the differences between network coding and pure routing for multiple independent unicast transmission sessions. The concept of network coding [1, 2] builds upon the fundamental observation that even in error-free communication networks, compared to routing only, applying coding operations at intermediate nodes as well as at the terminal nodes introduces extra flexibilities and new possibilities, which lead to new understandings in achieving high-rate, low-cost, or fault-tolerant data transmission.

Previous directions in network coding studies include, among others, characterizing the feasibility of coded transmissions [1, 2], investigation of the sufficiency of linear codes [3, 4], computation of coded flow topology [5, 6], code assignment [7], comparison with routing [8, 9], application in overlay and wireless networks [10, 11].

In previous work [9], we analyzed the benefit of network coding in improving throughput for single transmission sessions, including single unicast, single broadcast, and single multicast. We proved that the improvement is always bounded by a constant factor 2. However, in the presence of multiple concurrent sessions, little is known regarding the structure of a coded transmissions scheme, therefore the comparison with routing becomes much harder. As a special case, the scenario of multiple broadcast sessions is still tractable. Multiple *broadcast* sessions can be viewed as a group communication, and is therefore essentially equivalent to one broadcast session. In this paper, we present our ongoing study on the benefit of network coding for multiple *unicast* sessions. We consider different versions of the problem, including in directed or undirected networks, with

1

integral or fractional routing. The concepts of integral routing and fractional routing correspond to the coding and transfer of scalars and vectors respectively, in coding terminologies. In particular, half-integral routing corresponds to operations on 2-D vectors.

We show that the differences between network coding and routing depend crucially on the model of the problem, similar to the case of a single multicast session. In directed networks, again the throughput improvement due to network coding is unbounded, and may increase at the speed of $\Theta(|V|)$ as the network size grows. For undirected networks with integral routing, there still exist configurations that are feasible with network coding but infeasible with routing only. For undirected networks with fractional routing, we show that the potential of network coding to help increase throughput in a capacitied network is equivalent to the potential of network coding to increase bandwidth efficiency in an uncapacitied network. We conjecture that these benefits are non-existent.

## 2   Review of the single session cases

Before starting the main discussion on multiple unicast sessions, we first review previous studies of the single session cases.

Sanders *et al.* [7] showed a pattern of three-tier directed multicast networks, in which the ratio of maximum throughput with coding over that with routing, hereafter referred to as the *coding advantage*, grows as $\Theta(\log(|V|))$, and therefore has no constant upperbound.

In previous work [9], we examined parameters of an undirected communication network with fractional routing, including packing ($\pi$), throughput ($\chi$), strength ($\eta$), and connectivity ($\lambda$). *Packing* is equivalent to throughput with routing only; *throughput* refers to throughput with network coding; *strength* characterizes the partition-connectivity of the communication group; and *connectivity* refers to link connectivity among the communication group.

In the single unicast case, we show that $\pi = \chi = \eta = \lambda$, which implies the achievable unicast throughput is the same for network coding and routing. In the single broadcast case, we show that $\frac{1}{2}\lambda \leq \pi = \chi = \eta \leq \lambda$, which also implies network coding does not make a difference in achievable throughput. Similar techniques can be used to show that in directed or undirected integral cases, network coding does not lead to higher achievable throughput for single unicast or broadcast either. In the single multicast case, we show that $\frac{1}{2}\lambda \leq \pi \leq \chi \leq \eta \leq \lambda$, which implies throughput improvement due to network coding is bounded by a constant factor of 2. Furthermore, this bound is valid as long as half integral routing is allowed.

For the case of a single multicast in an undirected network with integral routing, it can be derived from Lau's [12] recent result $\lambda \leq 38\pi_I$ that $\chi_I/\pi_I \leq 38$. Therefore the coding advantage is also finitely bounded in this case.

Coding advantage values that have been observed in practice are at most slightly larger than 1.0 in the fractional case, and at most 2 in the integral case. Whether and how far the bound 2 in the fractional case and the bound 38 in the integral case can be improved are open problems.

# 3   The case of multiple unicast sessions

We use $G = (V, E)$ and $G = (V, A)$ to denote the network topology in the undirected and directed cases, respectively. The total number of independent unicast sessions is denoted as $k$. $S_i$ and $T_i$ are the sender and receiver of session $i$, for $i \in [1..k]$. Unlabeled links in graphical illustrations are assumed to have unit capacity.

In the single session cases, the feasibility of achievable coded throughput can be nicely characterized. A throughput demand $y$ is feasible if and only if a directed flow of rate $y$ can be set up from the sender to each receiver. In the multiple session cases, inter-session coding comes into play, and the problem becomes much harder. No such simple characterizations are known, even if all the sessions are unicast.

Without network coding, the multiple unicast routing problem becomes the multi-commodity flow problem (MCF). In the fractional model, MCF can be solved as linear optimization. In the integral model, MCF is equivalent to the general version of the edge-disjoint path problem, which is a well-known NP-hard problem in both directed and undirected settings.

## 3.1   Directed networks

Previous experiences in the multicast case show that the gap between network coding and routing is more evident in directed networks — the coding advantage is upper-bounded by 2 in undirected networks, while it may grow arbitrarily high in directed networks. Intuitively, one can orient any undirected network in the way that exactly matches the requirement of network coding, which may be adversary for routing, and therefore leads to a large gap between the two. In the multiple unicast case, the picture remains similar.



(a) Two unicast sessions in a directed network, each with throughput requirement 1.

(b) A coded transmission scheme that realizes these two unicast sessions.
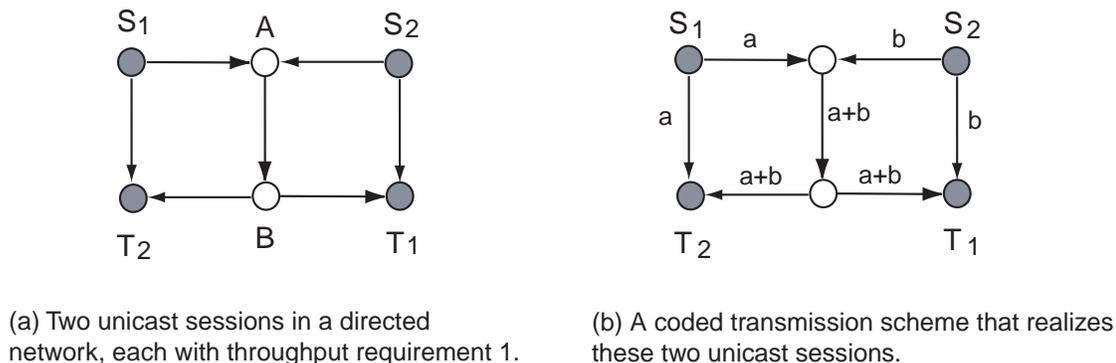
Figure 1: Coded unicasts in directed networks.

Fig. 1(a) shows an example scenario where two unicast sessions are to be set up in a directed network, with throughput requirement $(1, 1)$. Fig. 1(b) shows a coded transmission scheme that realizes both sessions without violating link capacity bounds or link directions. With routing only, the rate vector $(1, 1)$ is infeasible. Note that there is only one path connecting $S_1$ to $T_1$, $S_1 \to A \to B \to T_1$, and there is only one path connecting $S_2$ to $T_2$, $S_2 \to A \to B \to T_2$. These two paths share the unit capacity link $A \to B$, which becomes a bottleneck.

The example in Fig. 1 contains two unicast sessions, and shows a coding advantage of 2. Networks with larger number of sessions can be constructed according to a similar

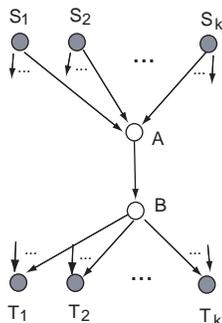pattern, such that larger coding advantages are exhibited.



Figure 2: A pattern of directed multiple unicasts that favors network coding. Each $S_i$ is connected to $A$ and all receivers except $T_i$; each $T_i$ is connected from $B$ and all senders except $S_i$.

As shown in Fig. 2, a directed network of the pattern has $2k$ unicast terminals, and two relay nodes $A$ and $B$. Every sender is connected to $A$, every receiver is connected from $B$, and $A$ is connected to $B$. Furthermore, every sender $S_i$ is connected to all receivers except the matching one, $T_i$ (which implies every receiver $T_i$ is connected from all senders except the matching one, $S_i$). With network coding, each receiver can send one bit information to $A$, at which point it is encoded with bits from other senders, and further relayed to $B$ and then each receiver. Besides, every sender will also send the same information to the $k - 1$ receivers it is directly connected to. As a result, every receiver $T_i$ is able to recover the 1 bit information from $S_i$. Without network coding, the total throughput of all the $k$ sessions is bounded by 1 bit. This can be verified by the fact that, removing the single link from $A$ to $B$ disconnects every sender-receiver pair in the network. In networks conforming to this pattern, network size $|V|$ is $2k + 2$, total throughput with coding is $k$, and total throughput with routing only is bounded by 1. We conclude that the coding advantage grows proportionally as $\Theta(|V|)$ in this case, with either integral or fractional routing.

## 3.2 Undirected networks with fractional routing

For the same configuration in Fig 1(a), if we make the network undirected by removing pre-fixed link directions, then both sessions are feasible if fractional routing is allowed. A half-integral two-commodity flow is shown in Fig. 3. The sufficiency of half-integer routing here is not just coincidence, as we shall soon discuss.
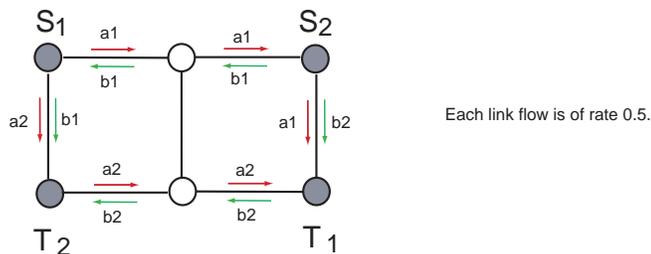


Figure 3: Uncoded fractional unicasts in undirected networks.

As to the question whether routing can always achieve the same throughput as network coding does, we first point out that in certain special scenarios, it is possible to obtain a confirmative answer by applying similar techniques used in the single session cases, *i.e.*, showing routing may achieve the cut value and hence a coded throughput value. By previous results in MCF studies, for the undirected fractional model, the cut condition is sufficient for flow feasibility in the following scenarios:

- When the number of sessions is one. In this case the cut condition also guarantees an integral optimal solution.

- When the number of sessions is two. In this case the cut condition guarantees a half-integral optimal solution.

- When all the senders or all the receivers reside on the same node. In this case an integral optimal solution always exists.

- When the network is planar, and all terminal nodes reside on the same face. This includes more specific cases such as when the network is a tree, when the network is a cycle, or when the network is planar and all terminal nodes lie on the outer infinite face. The cut condition guarantees half-integral optimal flows in this case.

However, in general the cut condition is not sufficient for MCF rate feasibility, even if flows are allowed to be arbitrarily fractional. A well-known counter example is given in Fig. 4. In this network, each of the four unicast sessions has a unit throughput requirement. we can verify that, the cut condition is satisfied. Yet no routing scheme can achieve this rate vector. To see the impossibility, note that the total network capacity $\sum_e c(e)$ is 6, while each sender is 2 hops away from its matching receiver and therefore total bandwidth requirement is at least $2 \times 4 = 8$.
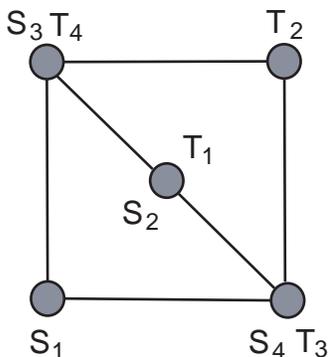


Figure 4: An example configuration where the cut criterion is not sufficient for multi-commodity flow feasibility. Throughput requirement is $(1, 1, 1, 1)$.

Although routing alone can not achieve the cut value in the above example , neither does network coding. A rate $\frac{3}{4}$ for each session is the best that can be achieved either with or without coding. We conjecture that it is actually always the case, *i.e.*, network coding may not improve achievable throughput in the undirected fractional network model. Below we present this conjecture as four equivalent propositions.

**Proposition 3.1.** For $k$ independent unicasts with a desired rate vector $y$, if the network is undirected and fractional routing is allowed, then $y$ is feasible with network coding if and only if it is feasible with routing only.

For the undirected fractional network model, there exists a fundamental connection between transmission rate and bandwidth efficiency. By modeling the fractional transmission problem as linear optimization and then apply the LP duality theorem, one may obtain the following observation: total throughput $y$ is feasible in a configuration $\alpha$ if and only if for any configuration $\beta$ in the throughput-equivalent class of $\alpha$, total network capacity is no less than $y$ times the minimum bandwidth consumption to achieve a unit throughput.

Here two configurations $\alpha$ and $\beta$ are called *throughput-equivalent* if one can be obtained from the other by applying the following two operations only: (a) *link expansion*, where a link is expanded into a path, with the same capacity at each hop as the original link, and (b) *path contraction*, which is the inverse operation of link expansion, where a non-branching path is contracted into a link, with the same bandwidth as the bottleneck bandwidth in the path. The intuition underlying this definition is that, a transmission scheme that serves a data communication configuration $\alpha$ may be straightforwardly modified to serve any other configuration that is throughput-equivalent to $\alpha$, to achieve the same throughput.

In the case of multiple unicast sessions without network coding, applying LP duality on the MCF LP leads to the following characterization of MCF rate feasibility, sometimes referred to as the Japanese Theorem [13, 14]: *a MCF rate vector $y$ is feasible if and only if for any distance function $x \in \mathbf{Z}_+^E$, $\sum_e c(e)x(e) \geq \sum_i d_i^x y_i$.* Here $d_i^x$ is the shortest path length between $S_i$ and $T_i$ under distance function $x$. Note that assigning an integer distance $x(e)$ for link $e$ corresponds to expanding a link $e$ to a $x(e)$-hop path, with the same capacity. Furthermore, the smallest bandwidth consumption to achieve one unit unicast throughput without coding is the shortest path length between the sender and the receiver.

We now know that the distance criterion in the Japanese Theorem implies MCF rate feasibility, which in turn implies coded unicast rate feasibility. Given these facts, Proposition 3.1 is equivalent to :

**Proposition 3.2.** A unicast rate vector $y$ is feasible with network coding, only if $\sum_e c(e)x(e) \geq \sum_i d_i^x y_i$, $\forall x \in \mathbf{Z}_+^E$.

Furthermore, since expanding a link to a path of the same capacity does not affect the rate feasibility, Proposition 3.2 in turn is equivalent to:

**Proposition 3.3.** If a unicast rate vector $y$ is feasible with network coding, then $\sum_e c(e) \geq \sum_i d_i y_i$, where $d_i$ is the number of hops on the shortest path between the sender and receiver nodes of session $i$.

Finally we show Proposition 3.3 is equivalent to Proposition 3.4 below. It is obvious that Proposition 3.4 implies Proposition 3.3, since bandwidth consumption is always upper-bounded by network capacity. On the other hand, if Proposition 3.3 holds, Proposition 3.4 must hold as well. Otherwise, take the routing flows with total size smaller than $\sum_i d_i y_i$ as the capacitied network, we obtain a counter example of Proposition 3.3.

**Proposition 3.4.** To achieve a unicast rate vector $y$ in a network with unlimited capacity at each link, the total bandwidth consumption is at least $\sum_i d_i y_i$.

## 3.3 Undirected networks with integral routing

In undirected networks with integral routing, network coding can make a difference in rate feasibility. Fig. 5 shows such an example, with two unicast sessions. Without coding, realizing the same two sessions requires half-integral routing.



(a) Two unicast sessions in an udirected network, each with throughput requirement 1.

(b) A coded transmission scheme that realizes these two unicast sessions, with integeral routing.
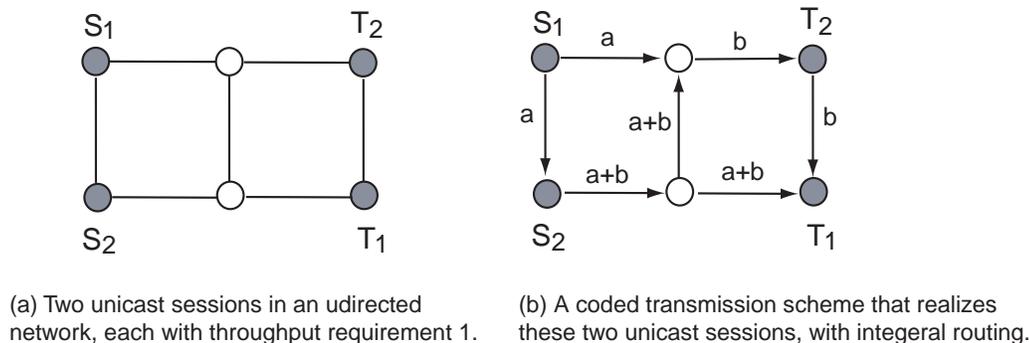
Figure 5: Coded integral unicasts in undirected networks.

The introduce of network coding does not narrow the gap between throughput and cut value. Fig. 6 depicts a well-known network pattern taken from the MCF literature, where either with or without network coding, only one unicast session can be supported at any time. However, the cut conditions for each session having a rate 1 are satisfied. Therefore with network coding, the gap between cut value and transmission rate remains at $\Theta(|V|)$. However, an appropriate characterization of the coding advantage in the integral case is still under investigation.
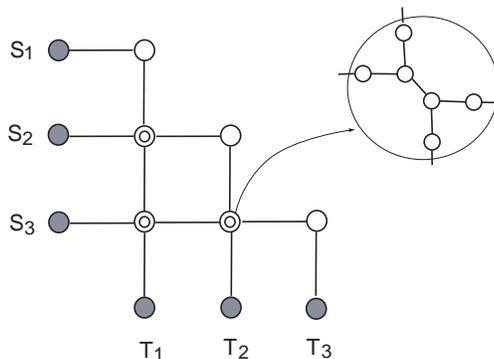


Figure 6: A network pattern where the gap between cut value and throughput grows as $\Theta(|V|)$.

## 4 Conclusions and future directions

The problem we study in this paper is whether network coding may lead to throughput improvement for multiple unicast sessions, and if so, how large the improvement can be. We show that for directed networks the coding advantage is not finitely bounded. In undirected networks with fractional routing, the question can be translated into whether network coding may save bandwidth consumption in a network with unlimited link capacities. The answer to this question is still to be investigated.

Our comparison of achievable throughput has been largely along the "network coding versus routing" direction, with a fixed integral or fractional routing model assumed. It is also interesting and important to compare achievable throughput with integral routing to that with fractional routing. For example, for a single multicast session or for multiple unicast sessions, only limited knowledge have been established regarding the differences among: maximum integral flow rate, maximum half-integral flow rate, and maximum fractional flow rate, either with or without network coding. The problem of computing maximum coded throughput for a single multicast in the undirected integral case is equivalent to the rooted group connectivity problem in combinatorics, and its complexity is also to be determined.

# References

[1] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[2] R. Koetter and M. Médard, "An Algebraic Approach to Network Coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, October 2003.

[3] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, pp. 371, 2003.

[4] M. Médard, M. Effros, D. Karger, and T. Ho, "Practical Network Coding," in *Proc. of the 41st Annual Allerton Conference on Communication, Control and Computing(Allerton)*, 2003.

[5] Z. Li, B. Li, D. Jiang, and L. C. Lau, "On Achieving Optimal End-to-end Throughput in Data Networks: Theoretical and Empirical Studies," Tech. Rep., Department of Electrical and Computer Engineering, University of Toronto, Feburary 2004.

[6] D. S. Lun, M. Médard, T. Ho, and R. Koetter, "Network Coding with a Cost Criterion," Tech. Rep., P-2584, Laboratory for Information and Decision Systems, MIT, April 2004.

[7] P. Sanders, S. Egner, and L. Tolhuizen, "Polynomial Time Algorithm for Network Information Flow," in *Proceedings of the 15th ACM Symposium on Parallelism in Algorithms and Architectures*, 2003.

[8] Y. Wu, P. A. Chou, and K. Jain, "A Comparison of Network Coding and Tree Packing," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, 2004.

[9] Z. Li and B. Li, "Network Coding in Undirected Networks," in *Proc. of the 38th Annual Conference on Information Sciences and Systems (CISS)*, 2004.

[10] Y. Zhu, J. Guo, and B. Li, "Multicast with Network Coding in Application-Layer Overlay Networks," *Journal on Selected Areas in Communications (JSAC), Special Issue on Service Overlay Networks*, vol. 22, no. 1, 2004.

[11] A. Ramamoorthy, J. Shi, and R. Wesel, "On the Capacity of Network Coding for Wireless Networks," in *Proc. of the 41st Annual Allerton Conference on Communication, Control and Computing(Allerton)*, 2003.

[12] L. C. Lau, "An approximate max-Steiner-tree-packing min-Steiner-cut theorem," in *Proc. of the 45th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2004.

[13] M. Iri, "On an Extension of the Maximum-flow Minimum-cut Theorem to Multicommodity Flows," *J. Operations Research Soc. of Japan*, vol. 13, no. 3, 1971.

[14] K. Onaga and O. Kakusho, "On Feasibility Conditions of Multicommodity Flows in Networks," *IEEE Transactions on Circuit Theory*, vol. 18, 1971.